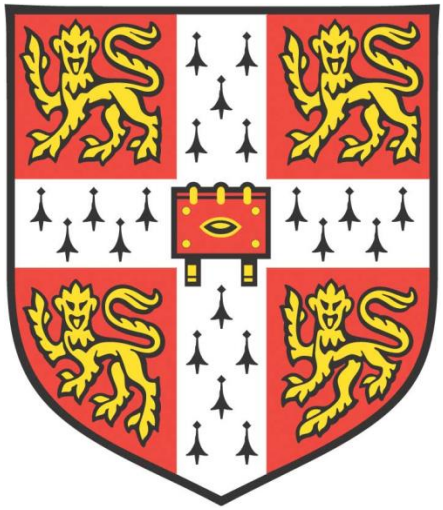
Representational dynamics across multiple timescales in human cortical networks



Tanya Wen
Christ's College

MRC Cognition and Brain Sciences Unit
School of Clinical Medicine
University of Cambridge

This dissertation is submitted for the
Doctor of Philosophy
September 2019

Preface

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

This thesis does not exceed the prescribed limit of 60,000 words (excluding tables, figures, appendices and references) as specified by the degree committee of the Faculties of Clinical Medicine and Clinical Veterinary Medicine.

The study outlined in Chapter 2 have been published in a peer-reviewed journal, and the studies outlined in Chapter 3 and Chapter 4 have been submitted to a preprint server. The references for these publications are as follows:

Wen, T., Duncan, J., & Mitchell, D. J. (2019). The time-course of component processes of selective attention. *NeuroImage*.

Wen, T., Duncan, J., & Mitchell, D. J. (2019). Representation of task episodes in human cortical networks. *bioRxiv*, 582858.

Wen, T., Mitchell, D. J., & Duncan, J. (2019). The functional convergence and heterogeneity of social, episodic, and self-referential thought in the default mode network. *bioRxiv*, 753509.

Some sentences in Chapter 1 and Chapter 5 have been taken from:

Wen, T., Mitchell, D. J., & Duncan, J. (2018). Response of the multiple-demand network during simple stimulus discriminations. *NeuroImage*, *177*, 79-87.

Wen, T., Liu, D. C., & Hsieh, S. (2018). Connectivity patterns in cognitive control networks predict naturalistic multitasking ability. *Neuropsychologia*, *114*, 195-202.

Tanya Wen

September 2019

Summary

Representational dynamics across multiple timescales in human cortical networks

Tanya Wen

Human cognition occurs at multiple timescales, including immediate processing of the ongoing experiences and slowly drifting higher-level thoughts. To understand how the brain selects and represents these various types of information to guide behavior, this thesis examined representational content within sensory regions, multiple demand (MD) network, and default mode network (DMN). Chapter 1 provides a background review of the current literature. It begins by reviewing experimental investigations of component visual processes that unfold over time. Next, the MD network is introduced as a collection of frontal and parietal regions involved in implementing cognitive control by assembling the required operations for task-relevant behavior. Finally, the DMN is introduced in the context of temporal processing hierarchies, with focus on its representation of situation models summarizing interactions among entities and the environment. The first experiment, presented in Chapter 2, used EEG/MEG to track multiple component processes of selective attention. Five distinct processing operations with different time-courses were quantified, including representation of visual display properties, target location, target identity, behavioral significance, and finally, possible reactivation of the attentional template. Chapter 3 used fMRI to examine neural representations of task episodes, which are temporally organized sequences of steps that occur within a given context. It was found that MD and visual regions showed sensitivity to the fine structure of the contents within a task. DMN regions showed gradual change throughout the entire task, with increased activation at the offset of the entire episode. Chapter 4 analyzed activation profiles of DMN regions using six diverse tasks to examine their functional convergence during social, episodic, and self-referential thought. Results supported proposals of separate subsystems, yet also suggest integration within the DMN. The final chapter, Chapter 5, provides an extended discussion of theoretical concepts related to the three experiments and proposes possible avenues for further research.

Acknowledgements

I would like to thank my supervisors Professor John Duncan and Dr Daniel Mitchell for taking me in as a PhD student. Our discussions on the workings of cognition have always been enjoyable and insightful. They not only gave me the freedom, but also the encouragement to pursue a variety of research topics. I am also grateful to Dr Caitlin Hitchcock for her mentorship, which has helped me better understand how to navigate the world. I have also benefited from many thoughtful conversations from past and present members of the MRC Cognition and Brain Sciences Unit and the wider University of Cambridge community.

I would also like to thank my previous mentors whom I have been in touch with for a long time and consider them friends. I would like to thank Professor Arthur Woodward who provided unconditional support during my applications to graduate school, and his teachings still inspire me in my academic career. I would also like to thank my undergraduate professor, Professor Shulan Hsieh, whose passion for and diligence in research has influenced me positively.

And to my family, for allowing me the freedom to pursue my scientific interests far away abroad, which meant less family time together.

Finally, I would like to thank the Taiwan Cambridge Scholarship from the Cambridge Commonwealth, European & International Trust for funding my PhD – I would not have been able to come here without it. I am also grateful for the financial support from the Medical Research Council PhD Studentship, as well as the Percy Lander Studentship in Preventive Medicine from Downing College.

Table of Contents

Chapter 1 Introduction

Human cognition unfolds at multiple timescales. It could occur as fast as a saccade to an item during visual search, or unfold slowly over minutes or even hours when carrying out a multi-step task. In any given moment, our interaction with the world consists of immediate, ongoing experiences, as well as higher level thoughts that drift in and out, that are not driven by immediate sensory input. It has been suggested that different parts of the brain represent different information occurring at different timescales to together give rise to "experience". In the following review, some of the salient findings concerning events and episodes at different timescales will be presented, as well as the techniques used to measure them

In the first section, the dynamics of attentional selection will be discussed. Early studies with electrophysiological recordings in behaving monkeys as well as ERPs/ERFs in humans have shown attentional selection to be remarkably efficient, in that it happens within a few hundred milliseconds, yet it consists of multiple component processes that occur successively in time. More recently, with the development of pattern analysis methods, it has been possible to track how representational dynamics evolve over time. The techniques and findings from high-temporal resolution recording data will be specifically reviewed

In the second section, a collection of frontal and parietal regions, known as the multiple demand (MD) network, will be reviewed. Activity in these regions has been found to be associated with many different kinds of task demands, and is believed to be involved in behavioral organization. The involvement of the MD network in goal-directed behavior is specifically reviewed, focusing on three aspects, including attentional selection, task difficulty, and task episodes. It is argued that these regions represent currently attended task-related information

In the third section, another network, the default mode network (DMN) will be reviewed. This set of regions highly overlap with areas in the brain involved in episodic memory. A recent theory of cortical information processing argues for a distributed topographical hierarchy of timescales, the DMN having the longest processing timescales. In this view, it is suited for representing high-level abstract features of the current event. The review will first focus on the involvement of the DMN in representing situation models and event boundaries. It will then raise the issue of the DMN's involvement in other cognitive

domains, including social cognition and resting-state activity. Better definition of the DMN, including division of subnetworks will then be discussed

Finally, a summary and overview of subsequent chapters, looking at various issues is presented.

1.1 Attentional Selection

In visual search, observers try to find a target object among distractors, in visual scenes where the location of the target is unknown. (Eimer, 2015) has summarized visual search to include at least four successive phases (Figure 1.1A). First is an initial preparatory phase, where the representation of the current search goal is activated. Once visual input has arrived, information of target-matching features is accumulated, which is then used to guide allocation of spatial attention to the target object, which then leads to encoding and recognition of the selected object.

For example, in the classic "Where's Wally" visual search game (Figure 1.1B), one must first have a mental representation of what the target (Wally) looks like before starting the search. Once the observer begins searching for Wally, objects with matching features (e.g., red and white stripes) will gain attentional weight, which will help guide attention towards those objects. Attention is later directed towards the location of those feature-matching objects before the object is recognized. Recognition of the object requires integration of features within the object, before a decision can be made on whether it is a target or non-target (Wally or a beach towel). As illustrated in this example, visual search is not a single process, but requires many stages of attentional processing that occur at successive time points.

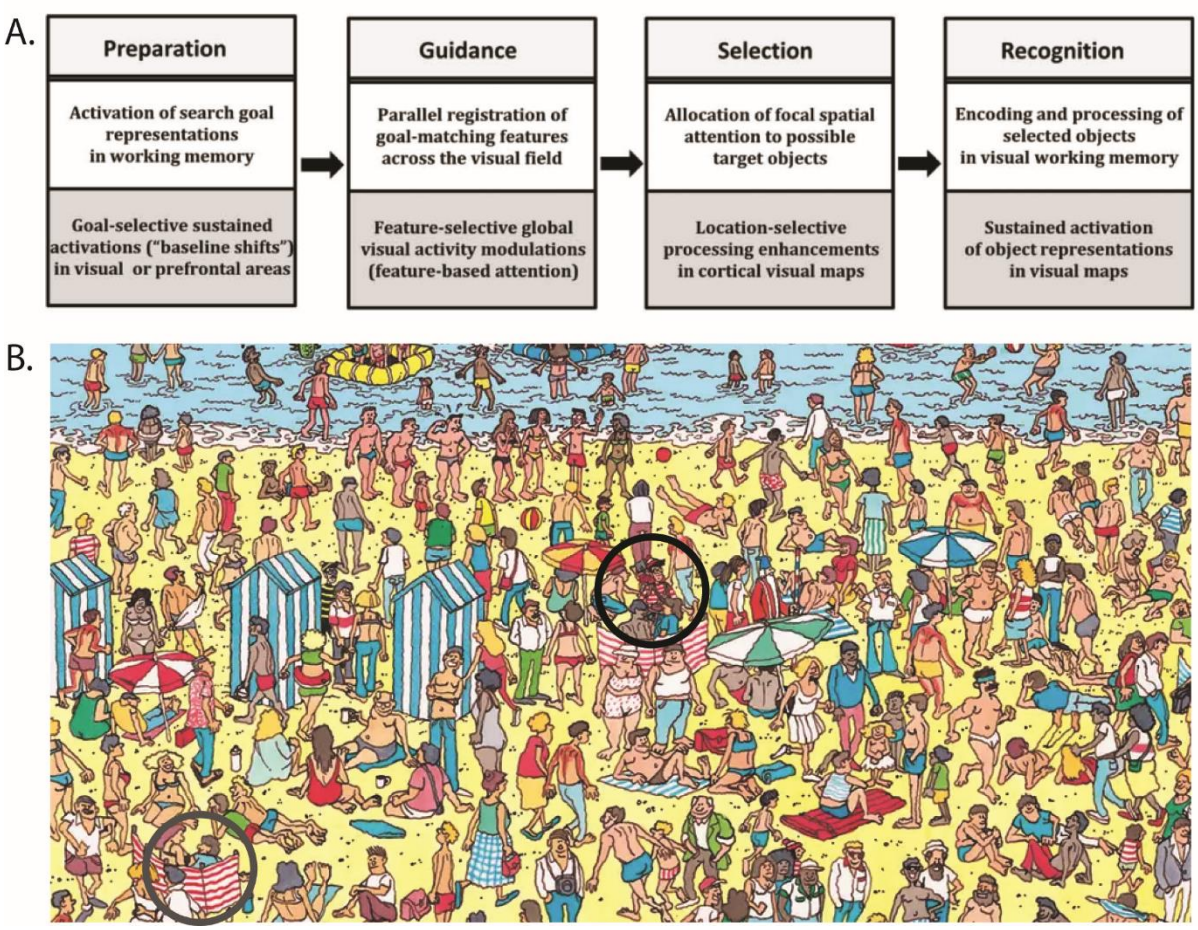


Figure 1.1. (A) Four phases of attentional control during visual search, including component processes of preparation, guidance, selection, and recognition, which emerge at successive points in time during a search process. The four functions are described at the cognitive level (white boxes) as well as their implementation at the neural level (gray boxes). Reproduced from Eimer (2015). (B) An example of a visual search task, where the target is presented among distractors in a visual scene. The black circle indicates the location of the target item, and the gray circle indicates a distractor item that shares similar features with the target (i.e., red and white stripes).

1.1.1 Preparatory attention

Before searching for a particular target object within a visual display, the observer first has to decide which object to search for. Next, a mental representation of the target is formed, known as the attentional template, and held "in mind". This preparation has been described by William James (1890) as "the anticipatory preparation from within of the ideational centers concerned with the objects to which attention is paid" (p 411). Preparatory attention may rely on the "central executive" of working memory that controls and manipulates its contents to match top-down goals, and working memory workspaces or "sketchpads" to maintain the

current attentional template (Baddeley and Hitch, 1974; Wolfe, 1994). These preparatory attentional templates are set up prior to visual search, which can help direct and guide attention to candidate targets later on during the search process.

An interesting question is how are such preparatory attentional templates implemented at the neural level? In a prospective coding framework, one could expect representation of attentional templates to resemble that of the anticipated stimulus. Therefore, one would expect that neural activity patterns in the visual cortex during the preparatory period would be similar to the patterns observed when the stimulus is perceptually processed (Awh and Jonides, 2001; D'Esposito, 2007; Stokes et al., 2009). This prediction has been investigated in single-unit recordings in monkeys, human fMRI, and human electrophysiology. In these experiments, neural activity is recorded during the preparatory phase (which is after a target-indicating cue, but before the onset of the visual search display). Early evidence for the neural correlates of attentional templates comes from Chelazzi et al. (1998). In the experiment, monkeys were given a delayed-match-to-sample task, where they were shown a picture of the target, followed by a delay period, before being presented with a visual search display (where they were required to make a saccade to the target if the target was present). During the delay period, the monkey simply fixated at the center of the screen with no further visual input, but the target had to be held in working memory in preparation for the search task. Interestingly, the delay period showed target-selective activity in inferotemporal regions, which was absent in a control experiment where the initial stimulus had no subsequent behavioral relevance.

Similar findings of increased baseline activity in regions sensitive to the target during the preparatory period have been reported in human fMRI studies (Battistoni et al., 2017). In these studies, participants were cued to attend for certain features, such as spatial location (e.g., Giesbrecht et al., 2006; Sylvester et al., 2009), motion (e.g., Chawla et al., 1999; Shibata et al., 2008), color (e.g., Chawla et al., 1999; Giesbrecht et al., 2006; Shibata et al., 2008), and object categories (e.g., Puri et al., 2009; Esterman and Yantis, 2010). During the delay period, activity has been found to be selectively enhanced in location-, motion-, color-, and object-sensitive regions depending on the cued feature. This has led to the view that maintenance of the attentional template is supported by similar neural mechanisms as seeing the target itself, known as the sensory-recruitment model of working memory (Awh and Jonides, 2001; Postle, 2006; D'Esposito, 2007; Serences et al., 2009). This line of research was extended with the development of multivariate pattern analysis (MVPA). It has been shown with fMRI that training a classifier on neural patterns elicited by stimulus processing can accurately predict the

contents of working memory during preparatory attention (Harrison and Tong, 2009; Serences et al., 2009; Stokes et al., 2009; Peelen and Kastner, 2011). For example, Stokes et al. (2009) had participants perform an independent task, where they viewed repetitions of two letters ("X" and "O") while detecting size deviants. The authors trained a classifier to discriminate the activity pattern evoked by these two letters, and were able to predict above chance which letter the participants were attending for in a separate attention task during the delay period following an auditory attentional cue. This cross-generalization between tasks demonstrates a shared neural code between the stimulus-evoked activity pattern and attentional template. The specificity of the top-down attentional bias during preparatory attention is thought to be reflected by the similarity of these two patterns, and has been supported by studies showing that the degree of target-specific activity and pattern specificity is correlated with behavioral performance (e.g., Giesbrecht et al., 2006; Stokes et al., 2009; Esterman and Yantis, 2010; Peelen and Kastner, 2011; Soon et al., 2013).

Electrophysiology studies using EEG and MEG have also found neural correlates of attentional templates. Early evidence comes from ERP studies showing a slow negative neural activity throughout the delay period of working memory tasks is elicited at posterior electrodes contralateral to the side of the target object during encoding. This is known as the contralateral delay activity (Vogel and Machizawa, 2004; Luria et al., 2016). One of the advantages of electrophysiological recordings is its high temporal resolution, which can be used to capture how activity and neural representations evolve over time (King and Dehaene, 2014). Some EEG studies have revealed that, for extended periods of time, information held in working memory cannot be decoded from a classifier trained to discriminate neural patterns elicited by visual perception using conventional methods (e.g., Wolff et al., 2015, 2017). Other studies have found decoding of the preparatory template during the pre-stimulus delay (e.g., Myers et al., 2015; Kok et al., 2017). However, in contrast to models that predict sustained pre-activation of the template, stimulus encoding was dynamic, and cross-generalization between stimulus encoding and template coding occurred only briefly around the time of stimulus onset. These results indicate that template activity patterns and stimulus activity patterns do cross-generalize, indicating a shared neural representation, but only weakly and transiently.

Recent findings from empirical data looking at single trial analysis of neural recordings suggest that spiking activity during the delay period is actually sparse, with brief bursts of activity with variable onset latency and durations sprinkled throughout the delay (Shafi et al., 2007; Stokes et al., 2013; Lundqvist et al., 2016a, 2018; Stokes and Spaak, 2016; Kucewicz et

al., 2017). It has been suggested that attentional templates may sometimes be stored in an "activity silent" passive form, such as changed synaptic weights (Lundqvist et al., 2010; Stokes, 2015). Furthermore, high-temporal resolution data show that the way information is reflected in brain activity is highly dynamic (Stokes et al., 2013; King and Dehaene, 2014). Time dependency can be expressed using cross-temporal generalization analysis, where discriminative patterns at each time point are statistically compared with discriminative patterns at every other time point (King and Dehaene, 2014). Stokes et al. (2013) showed, in a delayed paired-associate task, robust decoding of the attentional cue can be found in the monkey PFC when training and testing on the same time-points, but there was little cross-generalization between time-points. This suggests that preparatory activity travels through a continuous series of states rather than maintaining a fixed memory or anticipation state. Together, these results suggest that an attentional template is formed following a decision of what to look for; and high-temporal resolution data can reveal how neural patterns associated with the preparatory template evolve over time.

1.1.2 The successive processes of visual search

Our perception of the external environment is shaped by attentional selection according to our behavioral goals. After an attentional template is set up in preparation for the upcoming task, the search process begins after the visual search display is presented. Since the location of the target object is unknown in the preceding delay period, the selection of possible target objects is therefore based on the visual information available from the visual search display. As the brain has a limited processing capacity, we are aware of only a small part of sensory information at a given time in a cluttered visual world. Therefore, representations of visual inputs compete to gain dominance through a process known as biased competition (Duncan, 1996, 2006; Beck and Kastner, 2009). When visual inputs compete, a stronger representation for a particular visual object will be at the expense of other objects' representations. During visual search, if one is looking for a particular target, attention biases the selection of objects by giving each an "attentional weight" determined by their match to the attentional template that resembles a flexible description of the target stimulus (Duncan, 1980; Bundesen, 1990; Desimone and Duncan, 1995; Bundesen et al., 2015). These top-down mechanisms play a key role in resolving biased competition, such that attention can be directed to stimuli of high behavioral priority.

Given that the function of such top-down biases is to guide the allocation of attention during search for known targets at unknown locations, these biases are believed to operate in a

spatially global fashion across the visual field (Eimer, 2015). Information across the entire visual field is processed in parallel, and is selectively weighted for features that match the current search goal (Wolfe, 1994; Bundesen et al., 2005). These spatially global biases in feature-based attention have been shown in both fMRI and electrophysiological studies, such that attention to the target stimulus feature (e.g., color, motion, or object categories) showed increased visual cortical responses to spatially distant, ignored stimuli that shared the same feature (Treue and Martínez Trujillo, 1999; Saenz et al., 2002; Serences and Boynton, 2007; Peelen et al., 2009; Zhang and Luck, 2009). For example, in Peelen et al. (2009), participants were cued to detect bodies or cars within scenes presented at certain locations (and to ignore scenes presented in other locations). Using MVPA, the authors found significant decoding of objects that belonged to the target category even when they were presented outside the focus of attention; in contrast, objects belonging to irrelevant categories did not show above-chance decoding, even when they were presented inside the focus of attention. This finding provides evidence that top-down biases operate across the visual field independent of spatial attention. Furthermore, in the study, visual search displays were presented only for 130 ms, suggesting that feature-based attention occurs very rapidly, and very early in time. Zhang & Luck (2009) used high-temporal resolution EEG data to examine feature-based attention. With ERP analyses, the authors found that feature-based attention was able to influence feedforward sensory activity, within 100 ms of the onset of the visual display, even for stimuli presented at unattended locations. These findings suggest that top-down search goals influence visual search very early on, in a spatially global fashion, while features that match the search goals gain attentional weight.

This feature-based attention weighting can highlight the presence of potential target objects in the search display, which provides guidance signals for subsequent allocation of focal attention to particular objects (Wolfe, 1994; Bundesen et al., 2005; Gray, 2007; Eimer, 2015). Hopf et al. (2004) showed that feature-based attention initially operated in a location-independent manner, and that selection of features guided allocation of attention to the location of the target. In their experiment, EEG and MEG were recorded while participants performed a visual search task (Figure 1.2A; the target could be a red or green “C”, and in the example shown here is red). The target and lure were placed either on the left or right of the screen, surrounded by a cloud of distractor “C”s, in the same orientation as the target on target-side only, nontarget-side only, on both sides, or on neither side. Interestingly, the presence of task-relevant features led to a change in ERP/ERF activity around 140 ms after stimulus onset, and

this effect was independent of the location of the target (Figure 1.2B). This effect was followed by attentional allocation to the location of the target, indexed by the N2pc component, that began around 170 ms post-stimulus (Figure 1.2C). The N2pc is a well-known index of focusing attention to a target, and is characterized by a negativity in the posterior electrodes around 180-300 ms, contralateral to the location of the attended item (Luck and Hillyard, 1994; Luck et al., 1997, 2012). It has been suggested that the N2pc component is related to the process of biased competition between visual objects, enhancing perceptual processing of the target and suppressing information from distractors (Hopf et al., 2004; Luck et al., 2012). This has also been thought to be related to modulations of single-unit activity, where an initial on-discharge was produced by both target and nontarget stimulus, followed by latter suppression of the nontarget response (Chelazzi et al., 1993a, 1998; Duncan, 2006). These findings show that a dynamic process is involved as competition gradually resolves, generally favoring the target stimuli, resulting in a later modulation of visual activity.

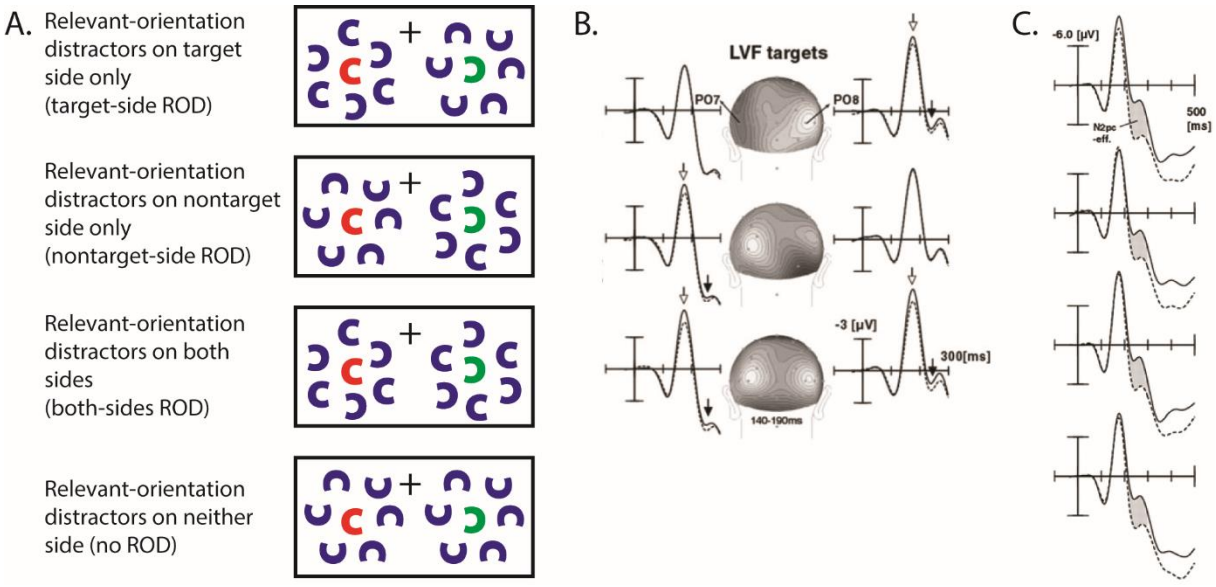


Figure 1.2. (A) Stimulus arrays used in the experiment of Hopf et al. (2004). Stimuli consisted of arrays of C-shaped items placed to the left and right of fixation, with one distinctively colored item in each field (red or green C). In this example, participants were required to attend to the red "C" and ignore the green distractor. The colors were randomized across trials. Distracter items (blue Cs) were placed in both visual fields surrounding the target and the lure. The orientation of the distracters was either left–right like the target (relevant-orientation distractor; ROD) or up–down (irrelevant-orientation distractors). The location of RODs and irrelevant-orientation distractors was varied to generate four different distractor distributions: (1) RODs in the target visual field and irrelevant-orientation distractors in the opposite field (target-side ROD); (2) RODs in the nontarget visual field and irrelevant-

orientation distractors in the target field (nontarget-side ROD); (3) RODs in both visual fields (both-sides ROD); and (4) irrelevant-orientation distractors in both visual fields with no RODs (no ROD). (B) ROD-related effects for left visual field targets. ERP waveforms elicited by target-side ROD (top), nontarget-side ROD (middle), and both-sides ROD (bottom, solid lines) trials superimposed onto waveforms of the no ROD condition (dashed lines) at selected electrode sites from the left parieto-occipital (PO7) and the right parieto-occipital (PO8) scalp. Similar trends were found in the right visual field (not shown here). (C) Target-related effects (N2pc effect). Average waveforms elicited by target items contralateral (solid line) and ipsilateral (dashed line) to electrode sites PO7/8. Data were collapsed over electrode sites to simplify presentation. The N2pc effect is highlighted by the gray areas between waveforms. Reproduced from Hopf et al. (2004).

After allocation of attention to the target location, further processing is required before the object is recognized. Object recognition takes place at a subsequent stage after selection, where the features of the selected objects are integrated, such that their identity becomes accessible (Xu and Chun, 2009; Carlson et al., 2013; Kiss et al., 2013; Eimer and Grubert, 2014; Eimer, 2015). In a study by Eimer & Grubert (2014), participants were asked to identify whether a target defined by a color-shape conjunction (e.g., blue square) was present among a display of four items. On no-competition trials, some displays contained the target object with three distractor items that shared neither of the two target features, while other displays contained a nontarget object that had one of the two target features (e.g., blue circles or red square, when the target was a blue square) and three distractors. On competition trials, the target and a partially-matching nontarget were present on the same display. Results showed that regardless of competition, both targets and partially-matching nontargets elicited reliable N2pc components. Interestingly, the sum of the two N2pc components elicited by color- and shape-matching nontargets was equal to the N2pc elicited by targets, until after 250 ms poststimulus, when the target N2pc became larger than its summed components. The superadditivity of the target N2pc suggested that selective attention processing changed from being feature-based to object-based, during which information is integrated across feature dimensions in later stages of processing. Integration implies that, as attention becomes focused on one object, all features of the object will be represented, not just those relevant to the task (Duncan, 1996; O'Craven et al., 1999; Schoenfeld et al., 2003). It has been proposed by Duncan (1996) that biased competition is integrated across different cortical regions, such that when attention is directed to one feature of an object, all features belonging to that object will become dominant in their respective cortical modules. For example, in Schoenfeld et al. (2003), when subjects attended

to one of two superimposed surfaces of moving dots, activity increased in the color-sensitive region of the fusiform gyrus when the surface moving in the attended direction displayed an irrelevant color (compared to when the dots were white). It has further been shown that following coding of exemplar-specific features, the brain can additionally categorize the objects on various levels of abstraction (e.g., animate or non-animate), and that the time it takes for the brain to discriminate such categories depends on the level of category abstraction (Carlson et al., 2013). Collectively, these findings suggest a temporal evolution from feature-coding to object-recognition to category-representation.

In summary, a simple visual search, that can happen within a few hundred milliseconds, involves multiple component processes. This begins with a flexible mental description of what to look for. During visual search, features matching the attentional template will gain attentional weight, guiding allocation of attention to potential targets. Finally, features are integrated and the objects can be recognized and categorized (Eimer and Grubert, 2014; Eimer, 2015). Many of these findings took advantage of high-temporal resolution electrophysiology recordings (e.g., single-cell, EEG, and MEG) that can capture dynamics at the millisecond timescale. Sensor-selection and source localization of EEG/MEG have suggested that a set of frontal and parietal regions play an important role in top-down modulation during working memory and visual processing (Corbetta and Shulman, 2002; Bressler et al., 2008; Simpson et al., 2011; Soon et al., 2013; Baldauf and Desimone, 2014; Wallis et al., 2015; Goddard et al., 2016). Simpson et al. (2011) showed a sequence of activation from early visual cortex to frontal and parietal regions during anticipatory spatial attention, corresponding to bottom-up extraction of cue meaning to top-down attentional deployment. Goddard et al. (2016) further used Granger causality analysis combined with MVPA, which showed feedforward and feedback flows of representational content between peri-occipital and peri-frontal areas during object recognition.

As mentioned above, it has been shown that a set of frontal and parietal areas have heavy involvement in attention. These regions are thought to be the source of the attention template (Duncan et al., 1997; Miller and Cohen, 2001), which generates top-down biasing signals to influence processing in other regions. Functional connectivity analysis has demonstrated significant Granger causality from the frontal and parietal cortex to visual occipital cortex during preparatory attention (Bressler et al., 2008). Finally, it has been shown that target detection not only elicits increased activity in sensory regions sensitive to the target features, but also in frontal and parietal areas (Corbetta and Shulman, 2002; Soon et al., 2013; Baldauf and Desimone, 2014; Finoia et al., 2015), suggesting these areas could be related to controlling

the locus of attention, or perform target-matching operations. The next section will focus more broadly on the functions of these frontal and parietal regions, collectively known as the multiple demand network.

1.2 The multiple demand network

It has been suggested that the frontal and parietal regions are particularly important for cognitive control (Miller and Cohen, 2001). In imaging studies, a characteristic pattern of frontoparietal activity, known as the "multiple demand" (MD) network (Duncan and Owen, 2000; Duncan, 2010), is produced by a wide range of different task demands. These regions have the ability to access and represent many different kinds of information, and adapt their coding in different contexts to code for specific information that is relevant to the current task demands (Duncan, 2001; Woolgar et al., 2011a).

A feature of human cognition is to be able to adapt and solve novel tasks. While some regions are specific to a particular cognitive function, e.g., sensory perception, movement, or language, regions involved in higher-level cognitive control need to be flexible and domain general (Duncan, 2006; Fedorenko et al., 2013). Several meta-analyses have consistently identified domain-general regions in frontal and parietal regions overlapping with the MD network (Duncan and Owen, 2000; Niendam et al., 2012; Yeo et al., 2015). In an early attempt in understanding the locus of cognitive control, Duncan (2006) (also see Duncan and Owen (2000)) reviewed 20 studies that manipulated five different cognitive demands: response conflict, task novelty, working memory load, working memory maintenance, and perceptual difficulty. Systematic comparisons of activations for the five cognitive demands showed a similar pattern of activation in the inferior frontal sulcus (IFS), anterior cingulate (ACC), anterior insula extending into frontal operculum (AI/FO), and intraparietal sulcus (IPS), shown in Figure 1.3A. This showed that while cognitive demands recruit only a set of specific parietal and prefrontal regions, the same network of regions are recruited in very diverse cognitive demands. Conceptually similar results were found by Dosenbach et al. (2006; Figure 1.3B), who conjointly analyzed data from mixed design experiments using 10 different tasks (varying in task rules, stimuli, sensory input, and response output) across a large number of subjects. Frontal and parietal regions, such as IPS and lateral prefrontal cortex (LPFC) were consistently activated at start-cue and error-related activity, suggesting a role in initiating trial-by-trial adaptive control, while ACC and AI/FO showed reliable start-cue and sustained responses in

nearly all tasks, suggesting a role in controlling goal-directed behavior through the stable maintenance of task sets.

More recent attempts to localize domain general activity found that this MD pattern exists at the individual subject level (Fedorenko et al., 2013; Figure 1.3C). Fedorenko et al. (2013) had participants perform 7 diverse tasks (including demands such as arithmetic, verbal/spatial working memory, and inhibition of irrelevant information), each with a difficult and easy condition, and identified regions that were responsive to cognitive demand as showing greater activity for difficult versus easy. Across all tasks, the same set of voxels showed consistent activity within the dorsolateral prefrontal cortex, extending along the inferior/middle frontal gyrus (IFG/MFG), and including a posterior-dorsal region close to the frontal eye field (pdLFC), parts of the anterior insular cortex (AI), pre-supplementary motor area and adjacent anterior cingulate cortex (pre-SMA/ACC), and intraparietal sulcus (IPS). The group analysis of Fedorenko et al. (2013) is now often used as an *a priori* MD template in a variety of studies on MD function (e.g., Crittenden et al., 2016; Muhle-Karbe et al., 2016; Tschentscher et al., 2017).

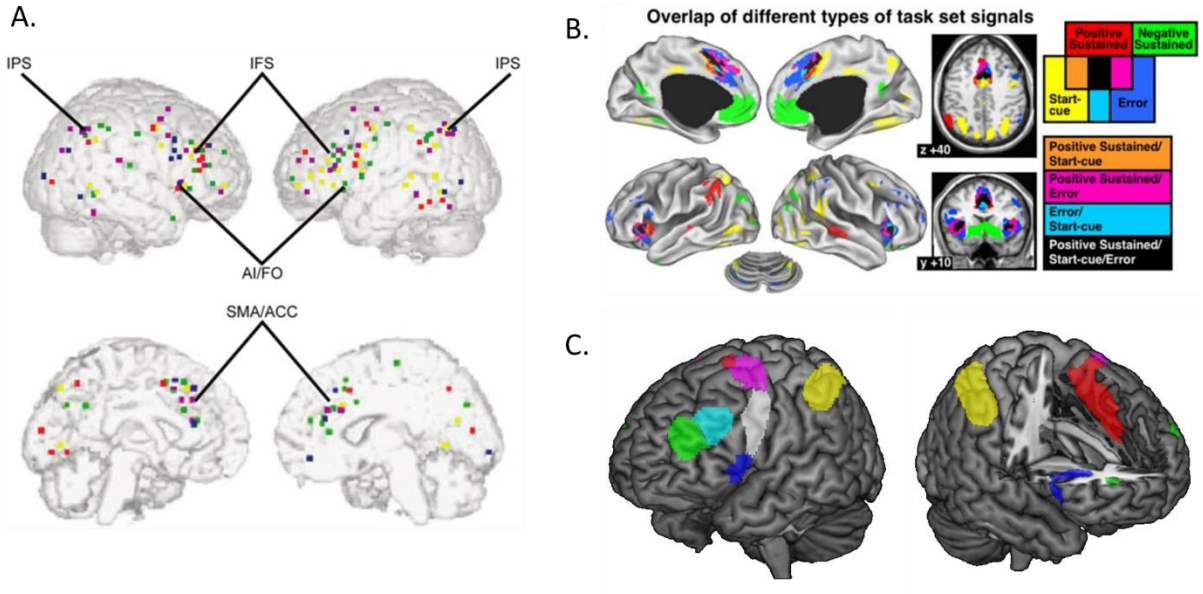


Figure 1.3 MD (multiple-demand) network in functional imaging. (A) Cortical activation foci from 20 studies examining response conflict (green), task novelty (purple), number of elements in working memory (yellow), working memory delay (red), and perceptual difficulty (blue). IFS: inferior frontal sulcus. AI/FO: anterior insula/frontal operculum. SMA/ACC: supplementary motor area/anterior cingulate. IPS: intraparietal sulcus. Reproduced from Duncan (2006). (B) Conjunction map of task-set signals showing regions that are active across different tasks. Reproduced from Dosenbach et al., 2006.

(C) Group-averaged multiple demand network from Fedorenko et al. (2013). ROIs taken from http://imaging.mrc-cbu.cam.ac.uk/imaging/MDsystem

1.2.1 Adaptive coding

In regions within the MD network, it has been shown that neurons may dynamically adjust their activity to carry information that is relevant to the current task demands. In a study by Freedman et al. (2001), the authors examined the neural activity in the monkey LPFC during an abstract categorization task. The authors used morphing software to generate stimuli that spanned two categories, "cats" and "dogs." Three species of cats and three breeds of dogs served as prototypes, with morphing used to produce all possible linear combinations between each of the six images (Figure 1.4A). By morphing prototype pairs in different proportions, the stimuli were varied continuously across each category boundary (i.e., more "cat-like" vs. more "dog-like"), or between two of the prototypes within each animal category. In the main task, monkeys were initially trained on a cat-dog categorization task, where on each trial they indicated with a lever whether an initial sample and subsequent test stimulus were from the same category. Among a large sample of responsive neurons in the LPFC, many were category selective, in that their activity differentiated between cats and dogs (Figure 1.4B), even when they were close to the decision boundary (e.g., cat:dog 60:40); but did not distinguish between morphs within a category (e.g., D1 and D3). It appears unlikely that outside the context of this particular task these neurons acted as cat-dog categorizers. To prove this, the authors retrained one of the monkeys on the categorization task with two new category boundaries which were orthogonal to the original boundary. This created three categories, which each centered around a prototype cat and a prototype dog (Figure 1.4A). After training, neural activity was discriminable among the three new categories, but no longer reflected the original categories, showing that coding preferences adapt to the context of the particular task.

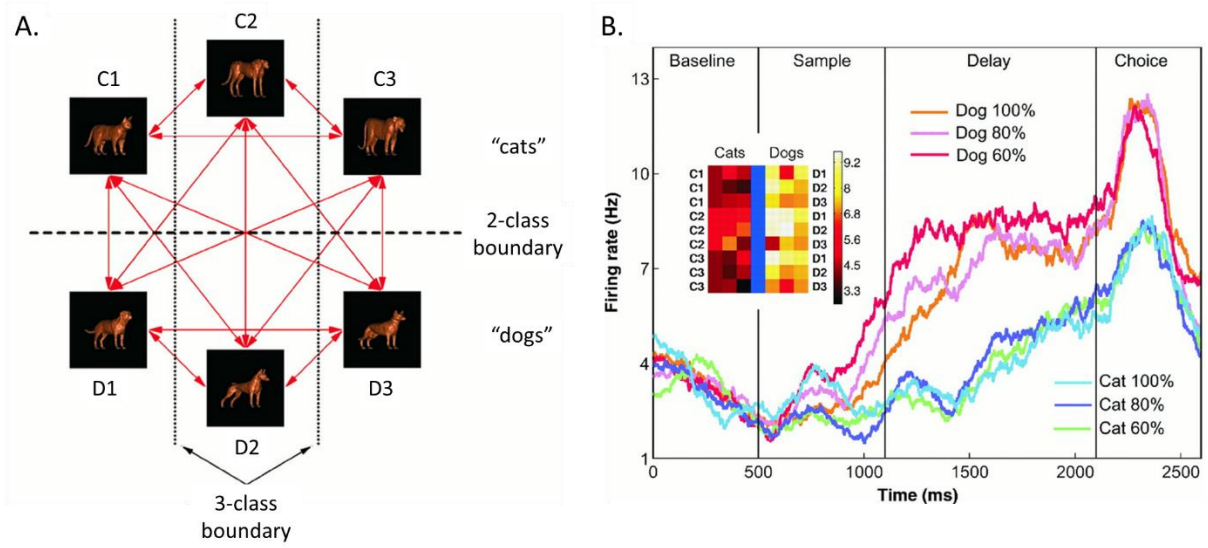


Figure 1.4 (A) Stimuli. Monkeys learned to categorize randomly generated "morphs" from the vast number of possible blends of six prototypes. Morphing software created blends between all prototype pairs (double-headed arrows), including within-category (for example, C1–C2) and between-category (for example, C1–D1) cases. (B) The average activity of a single neuron in response to stimuli at the six morph blends. This neuron responded strongly to all dogs, but poorly to all cats, irrespective of proximity to the category boundary. In the inset, the prototypes (C1, C2, C3, D1, D2, and D3) are represented in the outermost columns; each appears in three morph lines. A color scale indicates the activity level. Reproduced from Freedman et al. (2001).

Other studies have similarly shown that neurons in MD regions are strongly influenced by task context, with a selective focus on information of current behavioral relevance. In studies by Kadohisa et al. (2013) and Stokes et al. (2013), monkeys were trained to perform a delayed paired-associate task, where they learned pairs of cues and targets. During the choice display, either the target, a distractor, or a neutral stimulus could appear, and monkeys were required to make a saccade after stimulus offset to the location of the target and maintain fixation if the target was not present. Both studies showed that during the choice display, neurons in the frontal lobe initially discriminated the physical properties of the stimulus regardless of context, but later in time, coding transitioned to representing behavioral significance (target vs. non-target), irrespective of the stimulus. Similar results showing coding of behavioral categories have been found using human fMRI (Erez and Duncan, 2015). Other studies have shown that MD regions can code for several different types of task-relevant information, such as rules, stimuli, and motor responses (Li et al., 2007; Woolgar et al., 2011b, 2011a, 2015; Jackson et al., 2017), and

that the representation of information in MD increases in accordance to cognitive demand (Woolgar et al., 2011a, 2015).

The MD network is incredibly flexible and can change the contents of its representation on a trial-by-trial basis. Furthermore, the ability to flexibly combine different rules may be critical in explaining our ability to rapidly learn and implement new rules to novel situations, including learning novel task procedures from instructions (Dumontheil et al., 2011), even within a single trial (Cole et al., 2013a). At the simplest level, single rules link individual sensory stimuli to their appropriate behavioral responses; however, many tasks are complex and require the concurrent application of multiple rules. It has been found that the neural coding of multiple rules in the MD network is compositional, such that the neural activity patterns for multiple rules can be decomposed into the patterns of the constituent simple rules (Reverberi et al., 2012b; Garner and Dux, 2015). Cole et al. (2010) designed a cognitive paradigm where each trial involved three distinct rules that had to be integrated with each other to achieve accurate task performance. 12 task rules were permuted to create 64 unique tasks, and among those, 4 tasks were used for participants to learn each rule prior to scanning, while the rest were novel to participants. Within the LPFC, classifiers trained to identify abstract task rules based on practice task activity patterns were able to cross-generalize to novel tasks, suggesting that this region is involved in the adaptive transfer and composition of multiple rules (Cole et al., 2011, 2013b).

A growing number of functional connectivity studies have revealed functional networks underlying the brain's intrinsic functional architecture that matches task-driven activity patterns (Smith et al., 2009; Laird et al., 2011). Recent advances in brain connectivity have suggested that the frontoparietal network serves as a flexible hub for adaptive task control (Dosenbach et al., 2007; Cole et al., 2013b). The frontoparietal network has especially high global connectivity (Cole et al., 2010) and resting-state global connectivity in the LPFC has been found to be correlated with individual differences in cognitive control capacity and general fluid intelligence (Cole et al., 2012; Hearne et al., 2016). It has further been shown that this network rapidly shifts its functional connectivity with other networks when performing a variety of tasks, and that these connectivity patterns could be used to identify the current task (Cole et al., 2013b). This suggests that the frontoparietal network can access and influence other networks to assemble the neural components required for specific tasks. Together, these lines of evidence suggest a central role of the MD network in cognitive control and adaptive implementation of task demands.

1.2.2 MD activity, task difficulty, and behavioral organization

In line with the involvement of the MD network during cognitive demands, activity in the MD network had been found to increase with increases in many kinds of task difficulty (Duncan and Owen, 2000), such as with additional subgoals (e.g., Farooqui et al., 2012), greater working memory demand (e.g., Manoach et al., 1997), resisting strong competitors (e.g., Baldauf and Desimone, 2014), task switching (e.g., Wager and Smith, 2003), or a wide range of other task demands (e.g., Jovicich et al., 2001; Marois et al., 2004; Crittenden and Duncan, 2014; Woolgar et al., 2015). In fact, task difficulty has often been used as a manipulation to define MD regions (Fedorenko et al., 2013; Assem et al., 2017). Increased activity in more difficult conditions can also be accompanied by stronger information coding, shown by multivoxel pattern analysis (Woolgar et al., 2011b, 2011a, 2015).

However, in some cases, MD activity can be rather independent of task difficulty (Muller-Gass and Schröger, 2007; Cusack et al., 2010; Han and Marois, 2013; Dubis et al., 2016). For example, Cusack et al. (2010) contrasted hard and easy trials of a task in which participants had to detect a barely perceptible ripple in an oscillating dot field and found no neural activation differences between the two sensory difficulty levels, despite substantial differences in behavioral performance, and robust BOLD contrast to a different task manipulation (attention switching). Dubis et al. (2016) re-analyzed the tasks in Dosenbach et al. (2006) and showed no sustained activity in a purely perceptual task. They further conducted a new, difficult, perceptually-driven task, and found that signals in these regions are negligible or absent when a task is driven by perceptual information alone; however, MD regions were involved when the tasks required the use of abstracted representations beyond perceptual ones. Furthermore, they found that activations in these areas were not driven by task difficulty, measured objectively by reaction time and accuracy, and subjectively using a questionnaire rating.

In an important study, Han and Marois (2013) investigated activity in parts of the MD system during a task in which three letter targets were to be identified in a rapid stream of digit non-targets. In the baseline condition, the three letters occurred in immediate succession. To increase demand, the authors either inserted a non-target into the series of three targets, or reduced exposure duration. While activity in frontoparietal areas increased with the addition of a distractor, exposure duration had little effect. To interpret their findings, Han and Marois (2013) appealed to the distinction made by Norman and Bobrow (1975), between data-limited and resource-limited behavior. Norman and Bobrow (1975) proposed that, for any task, some

function (the performance-resource function or PRF) relates performance to investment of attentional resources. When this function is increasing, behavior is said to be resource-limited, and additional investment is repaid by improved performance. When the function asymptotes, further investment has no positive effect, and performance is said to be data-limited. In line with a link of MD activity to attentional investment, Han and Marois (2013) used these ideas of data- and resource-limitation to explain their findings. They proposed that, in their task, brief exposure duration created data limits, which could not be offset by increased frontoparietal recruitment, while adding a distractor introduced resource limits by calling for increased attentional focus.

Furthermore, while many studies in the literature show increasing MD activity with increasing task difficulty, there have also been studies that showed decreased MD activity (Bor et al., 2003), an inverted U-shape response (Callicott et al., 1999; Linden et al., 2003), or a plateau after a certain difficulty level (Todd and Marois, 2004; Marois and Ivanoff, 2005; Mitchell and Cusack, 2008). Callicott et al. (1999) and Linden et al. (2003) found that the frontal-parietal network initially showed increased activation with increased working memory load, but decreased in the highest load condition close or beyond the limit of capacity. In a particularly interesting study, Bor et al. (2003) compared the involvement of the MD network in encoding structured versus unstructured sequences into spatial working memory. They found that while reorganizing working memory contents into high level chunks decreased task difficulty and made the trials easier to remember, it recruited more MD activity than trials that did not allow chunking. This further suggests that MD is not responsive to difficulty *per se*, but recruitment of cognitive resources may be the driver of MD activation, especially when manipulations require behavioral organization beyond stimulus information.

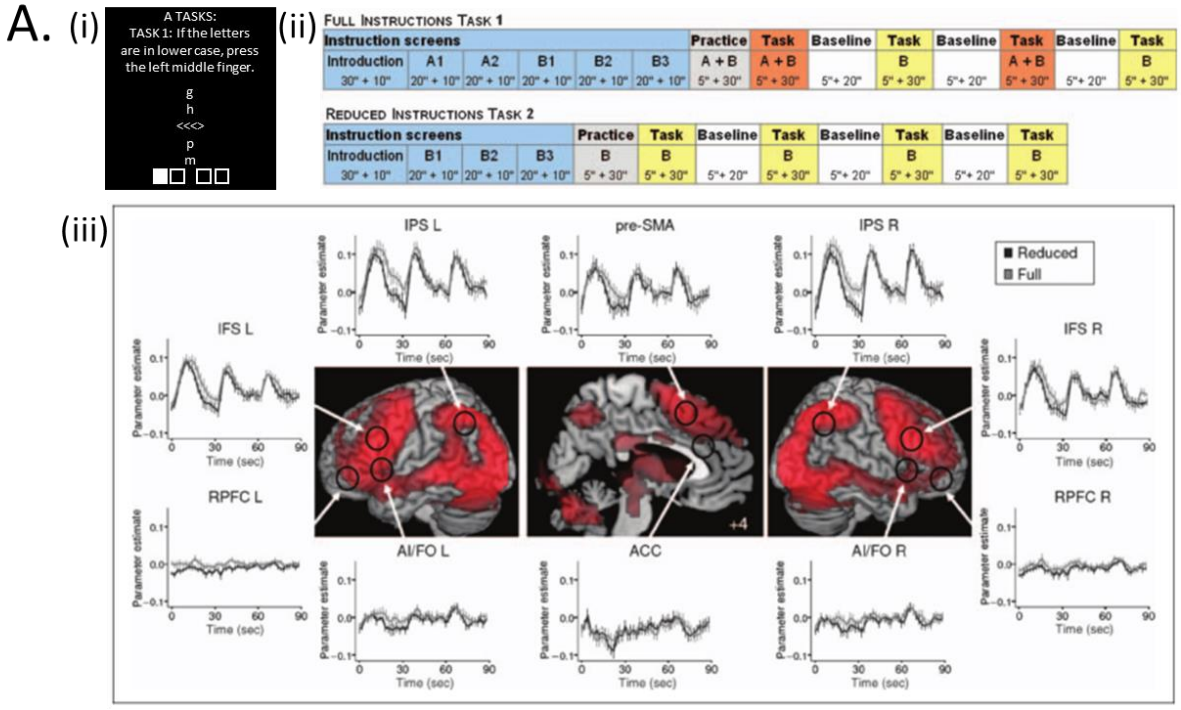
The above evidence suggests that the recruitment of the MD network plays a role in implementing top-down attentional control, optimally focusing processing for the requirements of a current task (Miller and Cohen, 2001; Duncan, 2013). While most work on cognitive control is concerned with isolated operations, such as attentional biases, target selection, or behavioral categorization, everyday behavior involves events that are situated in a larger context of a purposive mental episode. For example, in many everyday situations, such as planning a dinner (Penfield and Evans, 1935), making coffee (Cooper and Shallice, 2000), or solving logical puzzles (Lashley, 1951; Duncan et al., 2017), goals are achieved by assembling a series of subtasks, each separately defined and solved. It has been proposed that MD regions play a core role in defining and controlling these mental programs for complex cognition

(Kurby and Zacks, 2008; Duncan, 2010, 2013; Farooqui et al., 2012). In achieving a multi-step goal, each step requires selective focus on currently relevant information. Steps follow in rapid succession, requiring the MD network to rapidly change its processing focus and content (Duncan, 2010). As mentioned above, the coding of information in MD regions is highly flexible, and can rapidly reorganize the mental focus of attended information (Downar et al., 2000; Freedman et al., 2001). Furthermore, single-cell recordings have shown that neurons in the MD network can separate successive task steps within a cognitive episode (Sigala et al., 2008; Saga et al., 2011). For example, Sigala et al. (2008) compared the similarity of neural activity patterns defined by three task phases (cue, delay, and target), and found that activity patterns were approximately orthogonal between each phase, maximizing the discrimination of processing content of each step.

In fMRI, the MD network has been shown to be involved in task rule assembly, as each new rule is added throughout the task episode (Dumontheil et al., 2011). In Dumontheil et al. (2011), the authors constructed eight sets of tasks, and, for each task, instructions could define up to five rules (Figure 1.5Ai; e.g., if the letters are in lower case or if symbols are presented, press the left middle finger. If the letters are in uppercase and there is one filled square, press the direction that most arrows point to; otherwise, if there are zero or two vowels, do nothing). At the beginning of each task, a series of instruction screens were presented. Each task rule was described in a separate instruction screen, presented for 20 s with a delay of 10 s between each instruction and the next (Figure 1.5Aii). Timecourse analysis showed that MD regions exhibited a phasic response locked to each rule and its baseline response increased as each new task rule was assembled into the mental program (Figure 1.5Aiii).

Data from fMRI also show strong MD activity when a new task episode is created and when transitioning from one episode to the next. Increased MD activity is often seen at the boundary between perceived events (Sridharan et al., 2007; Kurby and Zacks, 2008). In a study investigating the involvement of the MD network in sequential behavior, Farooqui et al. (2012) designed series of hierarchical task episodes, for example participants were instructed to sequentially search for four target letters, such as “CATX” (Figure 1.5Bi), where completion of “CAT” completes the first subtask, and completion of “X” completes the entire episode. The results showed increased MD activity after each sub-goal is completed, with the greatest increase when the entire episode was completed (Figure 1.5Bii). Furthermore, in a second experiment, where the completion of the subtask was on step two (e.g., searching for targets “OHMY”), activation on step two was greater than step three in many MD regions. The authors

suggest that this increase in MD activity after completion of each step or subtask is a result of directing and revising the control representations of each step of the episode.



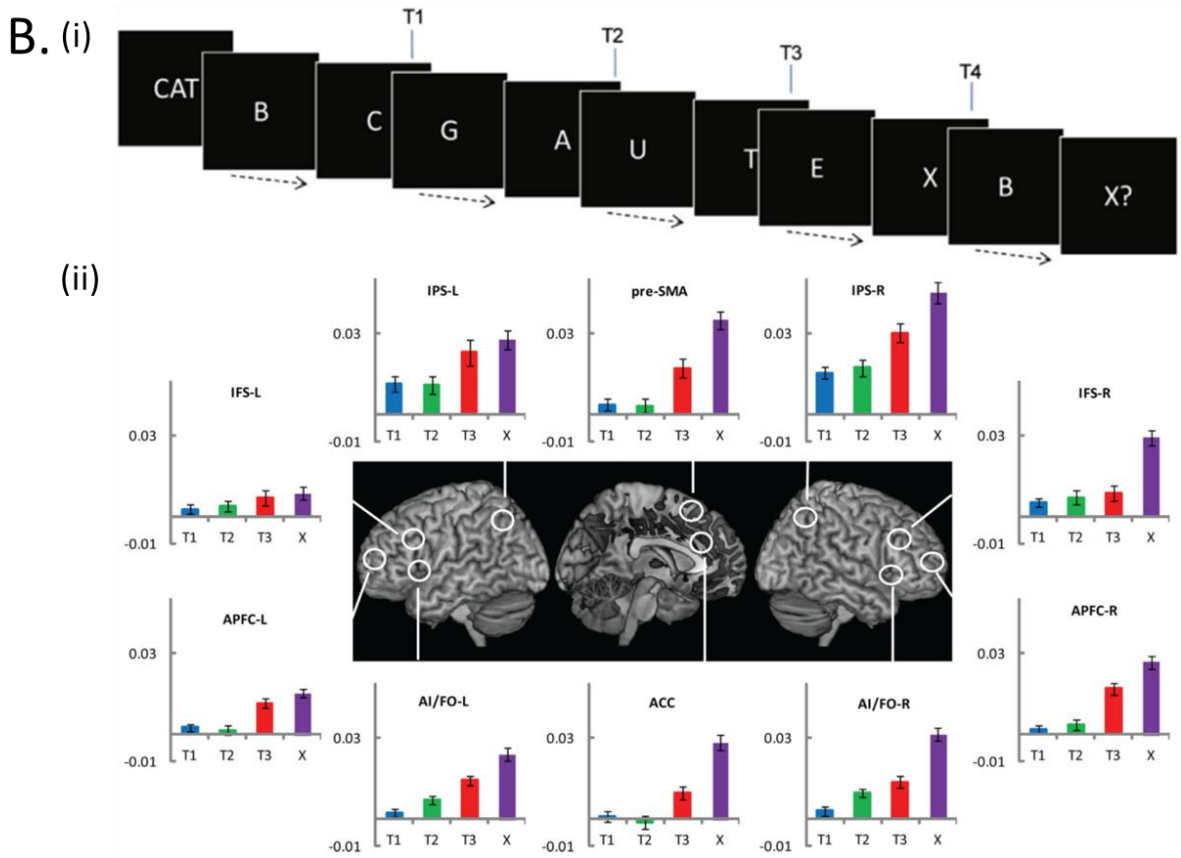


Figure 1.5 (A) (i) Example of an instruction screen for one of the eight tasks used in Dumontheil et al. (2011). For each task, instructions could define up to 5 task rules (e.g., if the letters are in lower case, press the left middle finger; if the display contains uppercase letters and there is one vowel, press the direction that most arrows point to). (ii) Illustration of task setup. At the beginning of each task, a series of instruction screens was presented. Each task rule was described in a separate instruction screen, presented for 20 s with a delay of 10 s between each instruction and the next. (iii) Timecourse analysis for each MD region of interest, showing phasic responses time-locked to instruction screens. Whole-brain render shows regions that were more active during instructions screens compared to baseline. Reproduced from Dumontheil et al. (2011). (B) (i) Structure of a typical trial in Farooqui et al. (2012). On each trial, participants monitored a sequence of letters, searching for a cued word (here "CAT"), and finally for the letter X. Targets were to be detected in the correct order. The first two targets (T1 and T2) completed subgoals at the lowest level (component letters of the first target word; level 1); the third target, T3, completed a subgoal at the next highest level (complete target word; level 2), while the fourth target, T4 (X), completed the whole goal of the task (level 3). Dotted arrows indicate the variable number of nontarget letters between initial cue, successive targets, and final probe. (ii) Comparison of phasic activity in response to various target events. Plots show activity index for each target. Reproduced from Farooqui et al. (2012).

In summary, the above review demonstrates the role of the MD network in a wide range of tasks. Coding of information in MD regions is highly flexible and can adaptively change in response to the current task demands. MD activity increases in accordance with cognitive demand, but can be independent of difficulty. Strong MD activity has been found during mental organization of behavior into chunks or mental programs, as well as completion of subtasks within an episode. It has been argued that the MD system plays a critical role in dividing complex problems into focused parts, and constructing component episodes (Duncan, 2010, 2013). However, the organization of sequences of events within a given context has also been associated with other regions involved in episodic memory (Ezzyat and Davachi, 2011; Eichenbaum, 2013; Hsieh et al., 2014; Cohn-Sheehy and Ranganath, 2017; Radvansky and Zacks, 2017), in particular in regions associated with the default mode network (DMN). The next section will review the involvement of DMN in representing events and situation models.

1.3 The default mode network

Representation of episodes has been widely studied in the context of episodic memory. Episodic memory contains representations of the record of spatial and temporal relationships between entities within a given context (Tulving, 1972; Ranganath and Ritchey, 2012), allowing people to re-experience the content and context of the events in the order they had previously occurred (Tulving, 2002). It has been suggested that different parts of the brain represent information at different timescales to together represent "experience". In everyday experience, we perceive a continuous stream of incoming sensory information, and this is integrated into longer timescale, meaningful representations. In an episode of "going to a formal dinner", one would need to "get dressed" and "eat dinner". While "getting dressed", subtasks could involve "put on a dress", "brush hair", and "wear a gown". As suggested previously, MD regions may segment episodes into smaller focused parts, selectively representing the currently relevant task (Duncan, 2010, 2013). However, while MD and sensory regions represent these immediate ongoing experiences, we have higher level thoughts (i.e., "going to a formal dinner") that may not be driven by our immediate sensory experience and occur over a longer timescale, spanning multiple events (Manning et al., 2015). Representation of these higher level episodes have been found to be related to regions in the DMN, which has been thought to integrate spatial and temporal information into situation models (Ranganath and Ritchey, 2012; Cohn-Sheehy and Ranganath, 2017). The DMN includes a set of functionally connected regions including the posterior medial cortex, medial prefrontal cortex, middle temporal gyrus, angular gyrus, and several other regions that are involved in episodic memory and retrieval (Raichle et al., 2001; Spreng and Grady, 2010; Rugg and Vilberg, 2013). An illustration of the DMN is shown in Figure 1.6A(a). The following sections will first review the role of the DMN in representing temporally extended episodes, and later discuss its broader involvement in other cognitive domains.

1.3.1 Long-timescale regions and event boundaries

According to some frameworks of hierarchical organization of functional brain networks that underlie goal-directed action, as one moves rostrally along the lateral frontal cortex, representations become increasingly more abstract (Badre, 2008; Badre and D'Esposito, 2009) and episodic (Koechlin et al., 2003; Koechlin and Summerfield, 2007). Dixon et al. (2014, 2017) have extended this framework to networks organized in a posterior-to-anterior gradient, starting from the sensorimotor network representing current stimulus and actions, to the anterior frontoparietal network (a network often coactivated with the DMN) representing distal goals. Hierarchical control contributes to tasks that require taking a series of actions in

time, where one must simultaneously manage a sequence of subgoals in the context of a superordinate goal (Badre and Nee, 2018). Temporally abstract episodic control has been found to produce more rostrocaudal and DMN engagement (Desrochers et al., 2015; Farooqui and Manly, 2018a).

These high level regions are not only involved in goal-directed tasks, but also represent temporally extended situations as we experience the world. Although we experience the world as a continuous sensory stream of information, people tend to segment these experiences into discrete events (Zacks et al., 2007; Radvansky and Zacks, 2011), and this segmentation can occur at different temporal granularities. Several studies have shown the existence of a hierarchy of temporal receptive windows in the human cortex, with early sensory regions representing short timescales, high-level sensory regions and MD regions representing intermediate timescales, and DMN regions representing long timescales (Hasson et al., 2008; Lerner et al., 2011; Honey et al., 2012; Chen et al., 2015). In one study, Chen et al. (2015) showed participants movie clips that were temporally scrambled at a fine timescale, scrambled coarsely, or intact. Inter-subject correlation was used to index the extent to which each region responded similarly across subjects to the continuous stimuli. They showed that voxels in primary sensory regions responded reliably across all subjects regardless of the degree to which the movie was scrambled, MD regions responded to both coarsely scrambled and intact movies, while DMN regions only responded reliably to the intact movie, suggesting a hierarchy of sensitivities to different temporal granularities (Figure 1.6A(b)). Baldassano et al. (2017) further showed that neural activity patterns during continuous experience are stable for a certain duration before rapidly shifting. The duration of pattern stability of each region corresponds to the timescale of event segmentation within the region, such that sensory regions encode short events, while DMN regions are relatively stable within temporally extended episodes (Figure 1.6B).

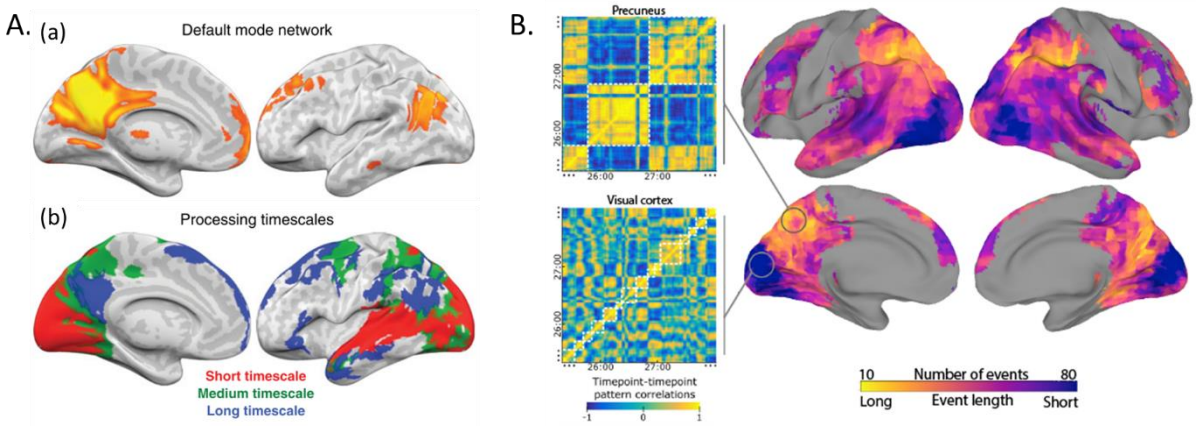


Figure 6. (A) (a) The DMN mapped by calculating functional connectivity (within-subject correlation) between the posterior parietal cortex (PCC) ROI and every other voxel in the brain. (b) Map of processing timescales across human cortex. Subjects viewed a movie clip either temporally scrambled at a fine timescale (<2 s), scrambled coarsely (7–22 s), or intact. For each clip, inter-subject correlation (ISC) was calculated at each voxel. Voxels were classified as "Long Timescale" if they responded reliably (i.e., above-threshold ISC) only when the movie was intact (blue); as "Medium Timescale" if they responded reliably during both the intact and coarsely scrambled movie (green); and as "Short Timescale" if they responded reliably in all 3 conditions (red). Reproduced from Chen et al. (2015). (B) The brain map shows the optimal estimated number of events identified by an event segmentation model by Baldassano et al. (2017) that identifies each rapid change in activity pattern as a new event. Results show a large number of short events in sensory cortex and a small number of long events in high-level cortex. For example, the time point correlation matrix for a region in the precuneus exhibited coarse blocks of correlated patterns, leading to model fits with a small number of events (white squares), while a region in visual cortex was best modeled with a larger number of short events. Reproduced from Baldassano et al. (2017).

Event boundaries occur when there is a perceived transition between a previous event and the next (Reynolds et al., 2007; Speer et al., 2007; Kurby and Zacks, 2008). It has been found that human annotations of meaningful event boundaries are associated with increased DMN activity (Speer et al., 2007; Yarkoni et al., 2008; Ben-Yakov and Dudai, 2011; Ezzyat and Davachi, 2011; Swallow et al., 2011) as well as pattern transitions in the DMN regions (Baldassano et al., 2017). Information about events is represented by mental models that capture the contents and structure of the events that people experience (Speer et al., 2007; Radvansky and Zacks, 2011), and DMN has been suggested to represent these high level abstract situations (Yarkoni et al., 2008; Baldassano et al., 2017; Cohn-Sheehy and Ranganath, 2017).

1.3.2 Posterior medial network and situation models

Context is often defined as slowly drifting information (i.e., information that persists over a relatively long timescale in the brain), which organizes our representations of co-occurring transient information (Howard and Kahana, 2002; Manning et al., 2014). A few examples of context could include information about the external environment, when an event occurred, internal thoughts, future plans, etc. (Smith and Vela, 2001). Many studies have found regions in the DMN are involved in contextual representation.

From earlier studies, it has been suggested that the medial temporal lobe is involved in representing events, and two subregions – the perirhinal cortex (PRC) and parahippocampal cortex (PHC) – have different functions. The PRC is thought to be involved in representing specific items, and the PHC is thought to represent information about the spatiotemporal context in which the items occurred (Diana et al., 2007; Ranganath, 2010a). It has later been found that these two regions are core components of two separate large-scale cortical networks. The PRC is a core component of an extended anterior medial system that is involved in representation of unique local entities. The PHC is a core component of the posterior medial system, which includes regions of the DMN, such as the posterior cingulate cortex (PCC), precuneus, angular gyrus, posterior lateral cortex (pIPL), and medial prefrontal cortex (MPFC), and has a central role in the representation of situation models (Ranganath & Ritchey, 2012; Reagh & Ranganath, 2018). A situation model is a mental representation that summarizes the interactions among entities in an environment, including place, temporal context, and social context, etc.

Regions that code for high level representations of situation models should be invariant of sensory modality, such that e.g., reading a book, hearing an audio narrative, seeing a movie, or recollection of the same episode should generate similar representations. Zadbood et al. (2017) and Baldassano et al. (2017) investigated how neural patterns associated with viewing television shows were encoded, recalled, and transferred to a group of naïve listeners. By comparing neural patterns across these three conditions, they found event-specific neural patterns in the DMN were able to cross-generalize across watching, recall, and listening to the same story. Interestingly, although both event durations and input modality in the three conditions were different, DMN exhibited shared neural patterns for each event regardless of condition. And although auditory cortex was active during both watching the movie and listening to the narrative, event structure did not cross-generalize in this region.

In addition to representing specific episodes, DMN regions have been shown to generalize across similar episodes, for example, various formal dinners at different colleges. Schemas are more abstract representations of situation models, involving organized semantic knowledge of stereotypical situations (Bartlett and Kintsch, 1932; Zwaan and Radvansky, 1998; Ghosh and Gilboa, 2014a). Several reviews have suggested that the MPFC is involved in schema representation (Preston and Eichenbaum, 2013; Robin and Moscovitch, 2017). Baldassano et al. (2018) analyzed fMRI data of human subjects that were presented with eight stories drawn from two different scripts (eating at a restaurant or going to the airport). Although each script shared similar structure (e.g., enter restaurant, being seated, ordering food, food

arriving), the stories varied widely in terms of their characters and storylines. They showed that patterns within the MPFC were sensitive to the overall script structure, abstracting away from the particular details of each story and activating a representation of the general type of situation being perceived.

These studies have shown that the DMN is involved in representing contextual episodes over a temporally extended period. Different regions in the DMN may represent the event scaffold at different levels of abstraction (context, situation model, and schema), together representing generalized temporal information of meaningful events (Cohn-Sheehy and Ranganath, 2017; Reagh and Ranganath, 2018). In addition to episodic representation, the DMN has been proposed to underlie a number of different processes (Buckner et al., 2008; Spreng et al., 2009; Andrews-Hanna, 2012), which will be discussed in the following section.

1.3.3 History, functional processes, and subnetworks

The DMN was initially identified by task-induced deactivations, where activity was found to be higher in a passive baseline condition compared to attention demanding external tasks (Shulman et al., 1997; Raichle et al., 2001). It was later introduced as a "task-negative" network by Fox et al. (2005). As the DMN is often identified during a passive baseline when no task was performed, it has been hypothesized to reflect spontaneous internal mentation (Buckner et al., 2008; Andrews-Hanna, 2012), such as mind-wondering (McGuire et al., 1996; Christoff et al., 2009). While the function of the DMN is still debated, more recent studies suggest that it is inaccurate to characterize it as "task-negative", as it has been shown to be actively engaged in a variety of tasks (Spreng and Grady, 2010; Spreng, 2012).

One aspect of DMN function is that it is related to self-projection (Buckner and Carroll, 2007). Many of these topics that have been traditionally studied independently seem to activate a similar network of regions related to the DMN. For example, studies of episodic memory have characterized a general recollection network (Rugg and Vilberg, 2013), which later has been found to share common neural substrates with imagination of future episodes (Addis et al., 2007; Schacter and Addis, 2007). In social cognition, a theory of mind localizer that had been developed to characterize regions involved in contemplating of other people's mental states has reliably found a similar network (Saxe and Kanwisher, 2003; Dodell-Feder et al., 2011). The most prominent regions within the DMN, the PCC and MPFC, have been found to be activated when engaged in self-relevant thought (Kelley et al., 2002; Benoit et al., 2010). Spreng et al. (2009) performed a systematic meta-analysis to quantify this common network of activation

among these various domains, and a conjunction analysis of "autobiographical memory", "navigation", "theory of mind", and "default mode" as keywords identified regions consistent with the DMN, with extensive functional overlap.

To dissociate the different components within the DMN, Andrews-Hanna et al. (2010b) performed a clustering analysis on intrinsic connectivity within the DMN, revealing three subnetworks – the dMPFC subsystem, the MTL subsystem, and the core hubs (Figure 1.7A). These three subsystems were similarly identified with resting-state ICA (Yeo et al., 2011). Task-based fMRI revealed a dissociation between the dMPFC subsystem and the MTL subsystem, which were preferably activated when participants made self-referential decisions about their present situation and when participants constructed events about their possible future, respectively, while the core hubs were activated in both self-relevant conditions (Andrews-Hanna et al., 2010b). In a later article, Andrews-Hanna (Andrews-Hanna, 2012) reviewed the three subsystems and their proposed functions (Figure 1.7B).

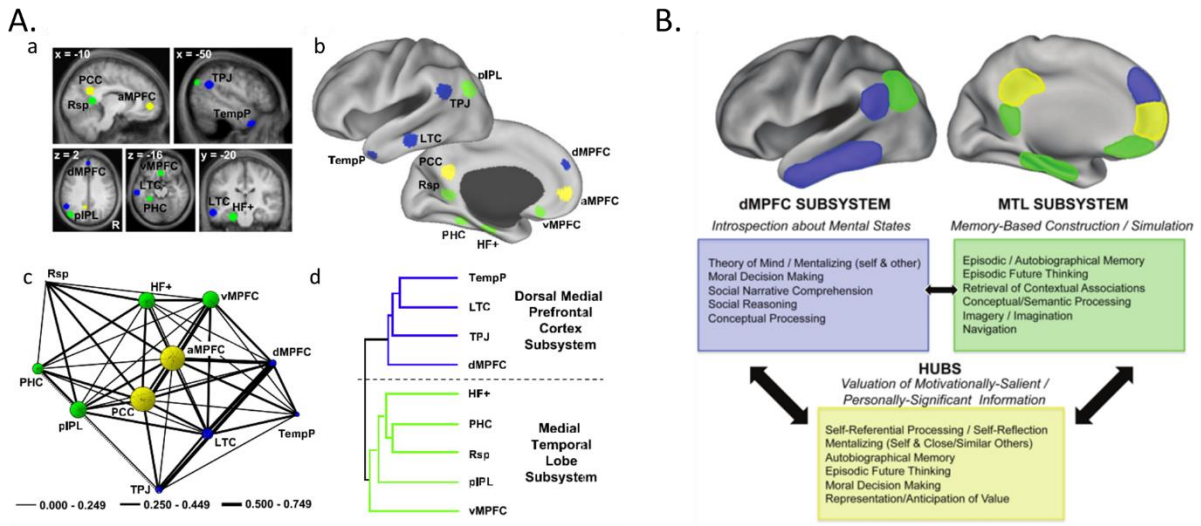


Figure 1.7 (A) (a)(b) Eleven a priori regions within the DMN in volume and surface space. (c) functional correlation strengths between the 11 regions, with the thickness of the lines representing the strength of the correlations between regions and the size of the circles representing how central a node is in the network. This revealed the aMPFC and PCC as core hubs of the DMN (d) Hierarchical clustering of the remaining regions, which grouped into two subsystems, the dMPFC and MTL subsystems. Reproduced from Andrews-Hanna et al. (2010). (B) Proposed functional-anatomic organization of the major default network components. A schematic drawing of the default network hubs (yellow) and subsystems (blue = dorsal medial prefrontal cortex subsystem; green = medial temporal lobe subsystem)

is highlighted along with each component's hypothesized functions and the tasks that frequently activate them. Reproduced from Andrews-Hanna (2012).

The dorsal medial prefrontal subsystem (dMPFC) consists of the temporoparietal junction (TPJ), lateral temporal cortex (LTC), and temporal pole (TempP) in addition to dMPFC. A rich body of work has suggested that the dMPFC subsystem may play an important role in introspecting about the mental states of social agents. These regions have been found to be consistently activated as a network when individuals reflect upon, evaluate, or appraise social information (Frith and Frith, 2006; Young et al., 2010; Dodell-Feder et al., 2011; Andrews-Hanna, 2012). For example, the dMPFC subsystem is activated when participants engage in interpersonal interactions (Rilling et al., 2004), view animated shapes to which intentions are imputed (Tavares et al., 2008; Isik et al., 2017), read social narratives (Fletcher et al., 1995), or reason about social problems (Van Overwalle, 2011) and moral dilemmas (Moll et al., 2005). Collectively, these findings suggest that the dMPFC subsystem plays a broad role in introspecting about self and other mental states.

The medial temporal lobe (MTL) subsystem consists of the ventromedial prefrontal cortex (vMPFC), posterior inferior parietal lobule (pIPL), retrosplenial cortex (Rsp), parahippocampal cortex (PHC), and hippocampal formation (HF). The MTL subsystem overlaps with the so-called core recollection network (Hayama et al., 2012; Rugg and Vilberg, 2013) or autobiographical memory network (Svoboda et al., 2006; Spreng and Grady, 2010). Studies looking at autobiographical memory (Cabeza and St Jacques, 2007; McDermott et al., 2009), episodic future thinking (Hassabis et al., 2007; Addis et al., 2009), retrieval of contextual associations (Bar, 2007), processing semantic and conceptual knowledge (Binder et al., 1999, 2009), and spatial navigation (Spiers and Maguire, 2006; Vann et al., 2009) have found the MTL subsystem to be engaged. It has therefore been suggested that a critical role of the MTL subsystem is to support memory-based construction and simulation (Schacter et al., 2008; Andrews-Hanna, 2012).

The core hubs consist of the posterior cingulate cortex (PCC) and anterior medial prefrontal cortex (aMPFC), and show the highest functional connectivity to all the other DMN regions (Andrews-Hanna et al., 2010b). The hubs of the default network activate across a diverse range of mnemonic, social, and emotional tasks that involve personally significant and other motivationally salient information (Olsson and Ochsner, 2008; Andrews-Hanna et al., 2010b).

These studies suggest that the DMN is involved in a wide range of tasks. While the precise functions of the DMN are not yet defined, and regions within the DMN could subserve different functions, evidence suggests that the DMN is involved in representing abstract high-level thoughts beyond those that emerge from the immediate environment.

1.4 Precis

In this thesis, different experiments are used to examine the temporal dynamics of information representation in sensory, MD, and DMN regions.

In Chapter 2, a combined EEG/MEG experiment is described that investigates the timecourses of component processes within a simple event that requires selection of a target in a visual display. With the advantage of millisecond precision recording, the experiment showed that multiple processes of selective visual attention can be robustly decoded from MEG/EEG signals. While the representation of a preparatory target template was largely undetectable prior to the choice display, the preparatory target template reemerged after presentation of a consistent non-target. During the visual display, attentional selection dynamically enhanced visual representations matching a target. Furthermore, the timing of directed attention to a target appeared earlier than explicit categorization of a single stimulus as a target.

Chapter 3 uses fMRI to examine the roles of the MD network and DMN during execution of larger task episodes, where a sequence of actions must be completed in the correct order to achieve a goal. The MD network and the visual cortex exhibited phasic responses to each task step, suggesting that they are sensitive to the fine structure of the episode. In contrast, DMN regions showed a phasic response predominantly to onset and offset of the entire episode. Beyond these phasic responses, gradually increasing activity across each task episode was seen throughout most of the brain. Representational similarity analysis revealed that both MD and DMN networks coded for step and item, while DMN additionally coded for episode.

Chapter 4 describes an experiment using six tasks to examine three different cognitive domains in which the DMN has been found to be involved. Theory of mind and moral dilemmas tasks were designed to engage the dMPFC subsystem; autobiographical memory and spatial navigation tasks were designed to engage the MTL subsystem; a self/other adjective judgement task was designed to engage the core hubs; finally, a resting baseline compared to a working memory task was carried out as to examine task-induced deactivation. While all six tasks activated the DMN, differences in activity patterns were found according to cognitive domain.

Finally, Chapter 5 discusses issues related to temporal organization, and possible roles of the MD and DMN networks.

Chapter 2 The time-course of component processes of selective attention

2.1 Abstract

Attentional selection shapes human perception, enhancing relevant information, according to behavioral goals. While many studies have investigated individual neural signatures of attention, here we used multivariate decoding of electrophysiological brain responses (MEG/EEG) to track and compare multiple component processes of selective attention. Auditory cues instructed participants to select a particular visual target, embedded within a subsequent stream of displays. Combining single and multi-item displays with different types of distractors allowed multiple aspects of information content to be decoded, distinguishing distinct components of attention, as the selection process evolved. Although the task required comparison of items to an attentional "template" held in memory, signals consistent with such a template were largely undetectable throughout the preparatory period but re-emerged after presentation of a non-target choice display. Choice displays evoked strong neural representation of multiple target features, evolving over different timescales. We quantified five distinct processing operations with different time-courses. First, visual properties of the stimulus were strongly represented. Second, the candidate target was rapidly identified and localized in multi-item displays, providing the earliest evidence of modulation by behavioral relevance. Third, the identity of the target continued to be enhanced, relative to distractors. Fourth, only later was the behavioral significance of the target explicitly represented in single-item displays. Finally, if the target was not identified and search was to be resumed, then an attentional template was weakly reactivated. The observation that an item's behavioral relevance directs attention in multi-item displays prior to explicit representation of target/non-target status in single-item displays is consistent with two-stage models of attention.

2.2 Introduction

Our perception of the world is constantly shaped by attentional selection, enhancing relevant over irrelevant information, to achieve our behavioral goals. Effective selection begins from a flexible description, often called the attentional template, of the object currently required

(Duncan and Humphreys, 1989; Bundesen, 1990). Much evidence suggests that attentional selection is then achieved through a process of biased, integrated competition across a broad sensorimotor network (Duncan et al., 1997). As objects in the visual input compete to dominate neural activity, the degree to which they match the attentional template determines their competitive advantage (Desimone and Duncan, 1995; Beck and Kastner, 2009).

Attention is often characterized as an emergent property of numerous neural mechanisms (Desimone & Duncan, 1995; Hopf et al. 2005), with different mechanisms dominating as successive stages of selection (Eimer, 2015). Therefore, while many studies have investigated the time-course of individual neural signatures of attention in humans and animal models, it is informative to compare multiple components of the selection process within the same paradigm. Recently, there has been much interest in the use of MEG/EEG for real-time decoding of cognitive representations in the human brain (Stokes et al., 2015). Here, we used simultaneous MEG/EEG to examine the time-course and content of different components of attentional selection. We combined single-item and multi-item search displays with different types of distractors to allow multiple aspects of information content to be decoded from the neural signal, distinguishing distinct components of attention as the selection process evolved.

The behavioral relevance of stimuli was manipulated by starting each trial with one of two auditory cues, indicating the relevant visual target object on this trial. Participants were then presented with a series of visual displays of 4 possible types: a 1-item display of the target (T), an inconsistent non-target (Ni; which was associated with the other cue and served as a target for other trials), a consistent non-target (Nc; which was never a target), or a 3-item display with all items presented simultaneously (see Figure 2.1 for an illustration). The use of inconsistent non-targets allowed representation of target status to be distinguished from representation of stimulus identity. The inclusion of 3-item displays allowed competitive representation of target location and target identity to be quantified under matched visual input. The use of consistent non-targets amongst a stream of choice displays allowed decoding of attentional template reactivation in preparation for a subsequent display. Participants made a button press whenever they detected a rare brightening of the target item. Requiring responses only for conjunctions of identity and brightening allowed response trials to be excluded from the analysis and attentional selection assessed on trials without an overt response. Using multivariate decoding analyses, we asked which component processes of attentional selection are visible in the MEG/EEG signal over time.

First, we examined representation of the attentional template. One possibility is that, when a cue indicates the relevant target object, some sustained signal will be set up in neurons selectively responsive to that object (Chelazzi et al., 1993b; Puri et al., 2009; Kok et al., 2013). fMRI decoding studies have shown cross-generalization between attentional templates and sensory responses to the corresponding objects (e.g., Stokes et al., 2009; Peelen and Kastner, 2011), supporting a tonic activation of visual representations. However, corresponding results tend to be weak or non-existent in electrophysiological recordings (Stokes et al., 2013; Myers et al., 2015; Wolff et al., 2015), and where they have been found, they may appear only very briefly prior to the target stimulus (Myers et al., 2015; Kok et al., 2017). Indirect measures of attentional templates, derived from ERP components, demonstrate that search templates are not continuously active but are transiently activated in preparation for each new search episode (Grubert and Eimer, 2018). Recently, it has been proposed that template storage may sometimes be "silent", perhaps encoded in changed synaptic weights rather than sustained firing (Stokes, 2015). To examine template coding, holding visual input constant, we analyzed data from the period between cue and displays, and during subsequent presentation of Nc stimuli.

Second, we were interested in the process of target selection itself. Comparing target and non-target stimuli shows strong differences both behaviorally and neurally (Duncan, 1980; Hebart et al., 2018). Attending to a relevant visual object produces strong, sustained activity across many brain regions (Desimone and Duncan, 1995; Sergent et al., 2005; Dehaene and Changeux, 2011), reflecting encoding of its multiple visual properties and implications for behavior (Wutz et al., 2018). In the presence of multiple stimuli, neural responses are initially divided amongst the competing sensory inputs and later become replaced by a wide-spread processing of the behaviorally critical target (Duncan et al., 1997; Kadohisa et al., 2013). On 1-item trials, we focused on the response to the T and Ni stimuli, to quantify the representation of object identity (e.g., face vs. house) regardless of status as target or non-target, as well as representation of behavioral category (T vs. Ni) regardless of object identity. On 3-item trials, we quantified the encoding of target location and target identity, to assess preferential processing of target features when multiple items compete for representation.

2.3 Methods

2.3.1 Participants

Eighteen participants (9 males, 9 females; age range: 18-30 years, mean = 24.4, SD = 3.8) took part in the study, recruited from the volunteer panel of the MRC Cognition and Brain Sciences Unit. Two additional participants were excluded from the analysis due to technical problems (one could not do the MRI; another was excluded due to an error in digitizing the EEG electrodes). EEG data for 4 participants were excluded from the MVPA analysis due to a technical issue (a test signal used during hardware checkup was not removed). All participants were neurologically healthy, right-handed, with normal hearing and normal or corrected-to-normal vision. Procedures were carried out in accordance with ethical approval obtained from the Cambridge Psychology Research Ethics Committee, and participants provided written, informed consent prior to the experiment.

2.3.2 Stimuli and Procedures

Participants performed two localizer tasks (auditory and visual) and an attention task (see Figure 2.1 for an illustration). Stimulus presentation was controlled using the Psychophysics Toolbox (Brainard, 1997) in Matlab 2014a (Mathworks, Natick, WA). Auditory stimuli were delivered through in-ear headphones compatible with the MEG recording. Visual stimuli were back-projected onto a screen placed in front of the participant, approximately 129 cm from the participant's eyes. Each stimulus image was approximately 20 cm wide (approximate visual angle 8.8°) on a gray background. Before the start of each task, participants were given training to familiarize them with the stimuli and task rules. If a false alarm was made during any of the trials during the recording, that trial was repeated at the end of the run.

2.3.2.1 Pattern Localizer Tasks

Auditory Localizer Task: This task was used to characterize multivariate activity patterns for high and low pitch tones used in the attention task. Participants heard a stream of intermixed high (1100 Hz) and low (220 Hz) pitch tones. On rare occasions (9% of the time), a frequency modulation would occur (modulator frequency = 20 Hz; modulation index = 2.5), and participants were instructed to press a button whenever they detected a distortion in a tone. There were 100 occurrences of each unmodulated tone and 10 occurrences of each modulated tone. The duration of each tone was 100 ms, with the beginning and ending 10 ms ramped. The inter-stimulus interval was jittered between 1000-1500 ms.

Visual Localizer Task: Similar to the auditory localizer task, this task was used to establish multivariate activity patterns for three visual stimuli (a face, a house, and a violin) used in the attention task. Participants were shown a stream of these images presented

sequentially in the center of the screen for 100 ms each, with an inter-stimulus interval jittered between 1500-2000 ms. Most image displays were semi-transparent (60% opaque) on a grey background; participants were asked to make a button press whenever they detected a brighter and higher contrast version of the image (100% opaque). There were 100 occurrences of each translucent image and 10 occurrences of each brightened image.

2.3.2.2 Attention Task

Figure 2.1 illustrates the stimuli used in the task, as well as the task structure. Before the start of this task, participants were trained to associate the two auditory tones with two of the three visual stimuli (the same used in the localizer tasks). This pairing resulted in the visual stimuli being categorized by behavioral relevance as targets (T: the visual stimulus paired with the current cue), inconsistent non-targets (Ni: the visual stimulus paired with the alternative cue), and consistent non-targets (Nc: never targets). All six possible mappings of two cues to three objects were counterbalanced across participants. The task was executed in runs of 90 trials. Each trial began with an auditory cue (for 100 ms), followed by a 1750 – 2250 ms fixation cross during which participants were instructed to prepare to attend for the target stimulus. Then a stream of three visual displays appeared one by one for 100 ms each, separated by 1500-2000 ms inter-stimulus intervals. Each display could be a 1-item display or a 3-item display with equal probability (order pseudorandomized, with the constraint that a 1-item display could not follow a 1-item display of the same type to minimize sensory adaptation effects). On 1-item displays, the stimulus was centered at fixation; 3-item displays contained all three visual stimuli, with the center of each stimulus 10° visual angle from fixation, arranged in an equilateral triangle with one above left, one above right, and one below. In 18 out of the 90 trials in each run, a single brightened stimulus, target or non-target, occurred pseudo-randomly in one of the 3 displays, with equal likelihood of appearing in each. For each cue type, brightenings affected T, Ni and Nc items once each on single item-trials, and twice each on 3-item displays, allowing one brightening for each of the six possible 3-item stimulus configurations. Participants were asked to attend to targets, pressing a button if they detected a brightened target (they could respond any time before the next stimulus), with no response for all other displays. Events with a brightened stimulus and/or button presses were later removed in the analysis, such that the results were not influenced by these events. The trial terminated if a button press was made, and participants were informed whether the response was a correct detection or a false alarm. A new trial began when the participant indicated with a button press that they were ready to continue. Otherwise, each of the 90 trials in each run had a full sequence

of 3 displays. At the end of each run, feedback informed participants of their accuracy through the run. To discourage false alarms and equalize the number of non-response trials across conditions, trials that contained a false alarm were repeated at the end of the run. The task was repeated over 5 runs (2 participants only completed 4 runs due to time constraints).

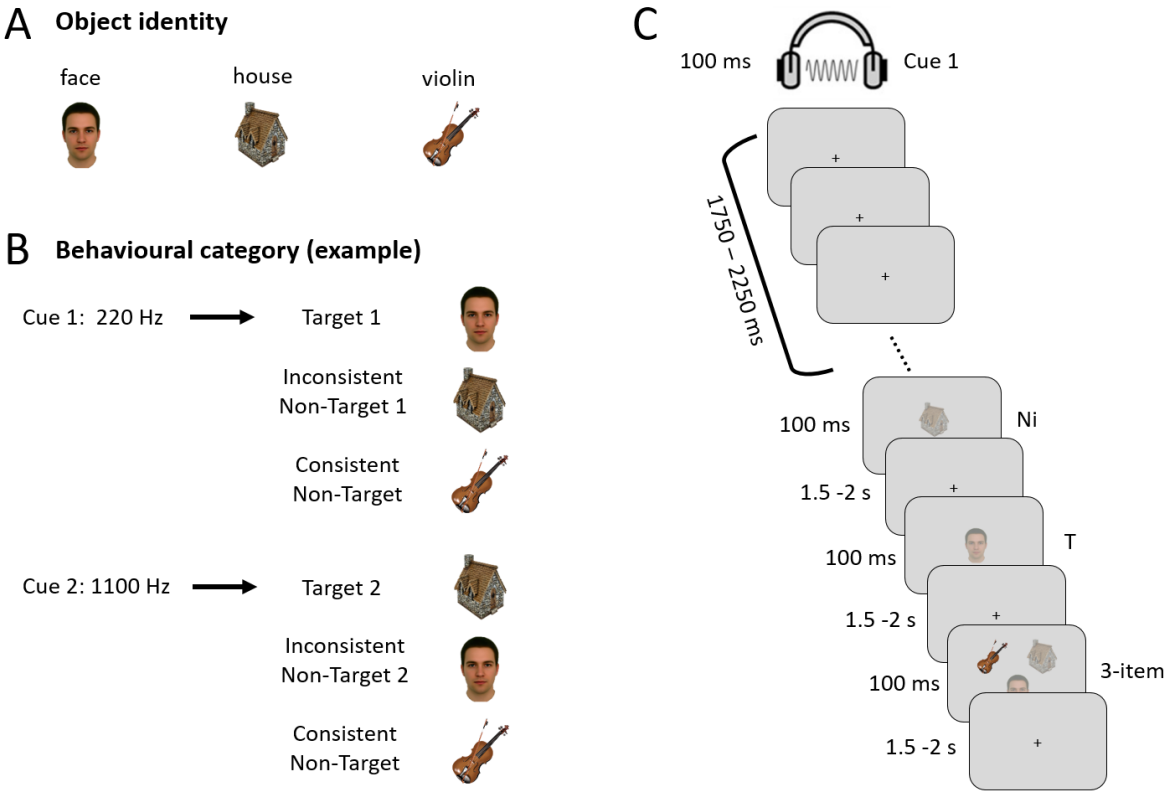


Figure 2.1. Stimuli and experimental paradigm. (A) The 3 objects used in the experiment. (B) An example of how the two auditory tones could be paired with the three objects. This results in two items that serve as targets (T) for one cue, and non-targets (Ni) for the other cue, and the third item serving as a consistent non-target (Nc). The pairings between the tones and the objects were counterbalanced across participants. (C) An example trial illustrating the experimental paradigm. At the beginning of each trial, an auditory cue indicated the target for that trial. After a delay, this was followed by three visual displays. Participants were asked to make an immediate button press if a brightening of the target stimulus was detected.

2.3.3 Data acquisition

2.3.3.1 Electroencephalography (EEG)

EEG data were collected from 70 Ag/AgCl electrodes mounted on an electrode cap (Easycap, Falk Minow Services, Herrsching-Breitbrunn, Germany) distributed according to the extended 10/20 system. Electrode impedances were kept below 5 kΩ. An electrode placed on the nose served as online reference while the ground electrode was placed on the right cheek. Vertical and horizontal eye movements were monitored using the electrooculograms (EOG) recorded using bipolar electrodes placed above and below the left eye and at the outer canthi of the eyes, respectively. Electrocardiography (ECG) was recorded using bipolar electrodes placed below the right collarbone and below the left ribcage. EEG data were sampled at 1000 Hz with a band-pass filter of 0.1–333 Hz. EEG and MEG data were acquired simultaneously.

2.3.3.2 Magnetoencephalography (MEG)

MEG data were acquired using a 306 channel (204 planar gradiometers and 102 magnetometers) Neuromag Vectorview system (Elekta AB, Stockholm) in a sound-attenuated and magnetically shielded room. Data were sampled at 1000 Hz with an online band-pass filter of 0.03–333 Hz. Five Head Position Indicator (HPI) coils were attached firmly to the EEG cap to track the head movements of the participant. The locations of the HPI coils as well as the EEG electrodes were recorded with a Polhemus 3D digitizer. We also measured three anatomical landmark points (nasion, left and right preauricular points) and additional points on the head to indicate head shape and enable matching to each individual's structural MRI scan.

2.3.3.3 Structural MRIs

High-resolution anatomical T1-weighted images were acquired for each participant (either after the MEG session or at least three days prior to the MEG session) in a 3T Siemens Prisma scanner, using a 3D MPRAGE sequence (192 axial slices, TR = 2250 ms, TI = 900 ms, TE = 2.99 ms, flip angle = 9°, field of view = 256 mm × 240 mm × 160 mm, 1 mm isotropic resolution). The coordinates of the nasion, left and right preauricular points in native space were hand-marked by the experimenter, and used in the coregistration of the EEG/MEG and MRI.

2.3.4 EEG and MEG data preprocessing

The raw data were visually inspected during recording for any bad channels, which were removed (EEG: 0 - 5 across subjects; MEG: 1 - 5 across subjects). The MEG data were de-noised using Maxfilter 2.2 (Elekta Neuromag, Helsinki), with the spherical harmonic model centered on a sphere fit to the digitized head points; default settings were used for the number of basis functions and the spatiotemporal extension (Taulu & Simola, 2006). Maxfilter detected additional bad channels using the first and last 900 data samples (default threshold), and signal

from all bad channels was removed and interpolated. Continuous movement compensation was applied at the same time.

Subsequent preprocessing used SPM12 (http://www.fil.ion.ucl.ac.uk/spm) and Matlab 2015a (The Mathworks Inc). Separately for EEG electrodes, magnetometers and gradiometers, independent component analysis (ICA), implemented using EEGLAB (Delorme & Makeig, 2004), was used to detect and remove components whose time-course correlated with EOG or ECG reference time-courses, and whose topography matched reference topographies associated with ocular or cardiac artefacts estimated from independent data acquired on the same system. ICA used the default infomax algorithm, with dimension reduction to 60 principal components. An independent component was removed if (1) it had the maximum absolute correlation with both a temporal and spatial reference, (2) these correlations were significant at $p < 0.05$, (3) the z-scored absolute correlations exceeded 2 for the spatial component, and 3 for the temporal component, and (4) it explained > 1.7% of total variance. For assessing temporal correlations only, ICA and reference time-courses were band-pass filtered between 0.1 - 25 Hz, and correlations were also repeated 1000 times with phase randomization of the reference time-course to ensure that the true maximum absolute correlation of eliminated components was greater than the 95th percentile of the null distribution. EEG data were then re-referenced to the average reference.

Data were band-pass filtered between 0.1 Hz and 40 Hz (zero-phase forward and reverse 5th order Butterworth filters with half-power cutoff frequencies). We note that although filtering enhances the signal-to-noise ratio of neural signals, it also spreads signal in time, distorting estimates of onset latencies. In this paper we focus on peak latencies, which are less sensitive to filtering (Luck, 2014; Grootswagers et al., 2016; van Driel et al., 2019). Data were epoched around the events of interest, time-locked to stimulus onset (from -100 ms to 1000 ms in the auditory localizer task; from -100 ms to 1500 ms in the visual localizer task; from -100 ms to 1750 ms for the cue and delay period of the main task, and -100 ms to 1500 ms for each of the visual stimulus presentations in the main task). Time points -100 ms to 0 ms served as baseline for baseline correction – the mean signal across this window was subtracted from each time point, per epoch. Epochs that contained flat segments or high threshold artifacts (peak-to-peak amplitude greater than 4000 fT for magnetometers, 400 fT/m for gradiometers, 120 μV for EEG, or 750 μV for EOG) were marked as bad trials and were rejected. In both localizer and attention tasks, any epoch that contained an auditory frequency distortion, a visual brightening, or a button press were additionally excluded from analyses. In the attention task, we also

removed all data from any trial with an error (false alarm or miss). The average number of epochs remaining for each condition is shown in Table 2.1.

Table 2.1. Mean number of epochs (and standard deviation across participants) per condition after artifact rejection

Localizer Tasks			
Auditory localizer			
Low tone	High tone		
80.8 (16.8)	81.0 (14.9)		
Visual localizer			
Face	House	Violin	
72.1 (13.6)	70.1 (14.9)	69.7 (15.4)	
Attention Task			
Preparatory phase			
Low tone	High tone		
147.6 (35.7)	147.4 (37.5)		
Stimulus processing phase			
Face	House	Violin	3-item
138.4 (36.4)	134.2 (34.2)	139.5 (34.3)	406.1 (110.2)
Target (T)	Inconsistent target (Ni)	Non-Consistent target (Nc)	Non-

138.9	137.9	135.1
(35.3)	(35.0)	(34.8)

2.3.5 Source Localization

For each participant, a cortical mesh was created from the individual's structural MRI, with a mesh resolution of ~4000 vertices per hemisphere. The EEG/MEG and MRI were coregistered based on the three anatomical fiducial points and an additional ~200 digitized points on the scalp. Forward models were computed for EEG data using a three-shell boundary element model (BEM) and for MEG data using a single-shell BEM. The forward model was inverted using minimum norm estimation (MNE) to compute source estimates for each experimental condition.

Due to the limited spatial resolution limits of EEG/MEG, we chose three *a priori* spatially distinct bilateral ROIs (Figure 2.2C). Early visual cortex and lateral prefrontal cortex (LPFC) were used to test representation in relevant sensory and cognitive control areas. An additional auditory cortex ROI was used both to measure cue decoding, and in other analyses to test for signal leakage. Auditory and primary visual cortex ROIs were taken from the SPM Anatomy toolbox (Eickhoff et al., 2005), containing 350 and 523 vertices. The LPFC ROI was taken from Fedorenko et al. (2013) (http://imaging.mrc-cbu.cam.ac.uk/imaging/MDsystem), combining the anterior, middle, and posterior middle frontal gyri, spanning 461 vertices.

We chose V1 as the visual ROI to keep the three regions as far apart as possible and thus minimize signal leakage between them. Since higher visual regions are specialized for object-level processing, and can contain template-like signals (Stokes et al., 2009), we subsequently examined a broad extrastriate visual cortex (ESV) ROI from the Fedorenko et al. (2013) template, which encompasses object, face, and scene processing regions. In all cases, results were very similar, reflecting the low spatial resolution of MEG. Here we report the results of V1, but the results from ESV can be found in Appendix A. Figures A.1~A.6.

2.3.6 Multivariate Pattern Analysis (MVPA)

Multivariate pattern analyses were performed using the Matlab interface of LIBSVM (Chang and Lin, 2011). We used a linear support vector machine (SVM), with default

parameters. For each analysis, we performed decoding in sensor space as well as in source space using data from the three ROIs. For sensor space decoding, we combined data from good EEG and MEG (gradiometers and magnetometers) channels. Each individual time point was standardized (z-scored across channels) before entering the classifier. For source space decoding, each participant's cortical mesh was transformed into MNI space, and estimated source activity at each vertex within the ROIs was extracted to serve as a feature in the classifier.

In both sensor and source space MVPA analyses, we trained and tested using spatiotemporal patterns extracted from a sliding time window of 32 ms, in 4 ms steps. Training and testing were performed on every combination of time windows, resulting in a cross-temporal generalization matrix of classification accuracies (King and Dehaene, 2014), with the diagonal representing the performance of classifiers trained and tested on the same time window. The classification accuracy matrix was then slightly smoothed using a sliding 32 ms square averaging window. For analyses involving within-task decoding, the data were split into five folds (with one fold containing every 5^{th} trial chronologically), iteratively trained on individual trials from four of the folds and tested on the remaining fold by applying the SVM to the remaining trials individually. In cross-task decoding, a classifier was trained on all relevant epochs from one task and tested on all relevant epochs from another task.

Classification accuracies were compared against chance (50%) with one-tailed t-tests. Multiple comparisons were accounted for using Threshold Free Cluster Enhancement (TFCE), with height and extent exponents of 2 and 2/3 respectively, and Family-Wise Error controlled by comparing the statistic at each time point to the 95^{th} percentile of the maximal statistic across a null distribution of 1000 permutations with random sign flipping (Smith and Nichols, 2009). TFCE was performed in the same way across the time × time decoding matrices and along the matched-time diagonals. The figures were plotted according to the last time bin in the sliding window (Grootswagers et al., 2017). For decoding of 1-item behavioral category, epochs that were preceded by a T or Ni were excluded, to ensure that behavioral category was balanced in the baseline period.

2.4 Results

2.4.1 Behavioral results

Behavioral performance was consistently high (auditory localizer task – hits: mean = 99.0%; false alarms: mean = 0.8%; visual localizer task – hits: mean = 98.9%; false alarms: 1.0%; attention task – hits: mean = 98.3%; false alarms: mean = 0.8%).

2.4.2 Coding of the attentional cue/attentional template during the preparatory phase

Source localization of the response to the cue at representative time points is shown in Figure 2.2A. We first looked for decoding of the specific attentional cue during the preparatory phase of the attention task, defined as starting from cue onset but before the first visual stimulus, and compared this with decoding in the auditory localizer task. We asked whether preparing for a target enhances cue decoding. Here, we subsampled the trials in the attention task to match the minimum number of trials in the auditory localizer for each participant, keeping the first n trials, to ensure comparable signal-to-noise ratio across the three decoding analyses. Cue/stimulus decoding as a function of time from auditory stimulus onset is shown in Figure 2.2B, D. Curves on the left show training and testing on matched time-points. Matrices on the right show generalization of patterns across all pairs of training and testing time windows.

Across the whole sensor space (Figure 2.2B), significant discrimination between the two auditory stimuli/cues emerged shortly after the presentation of the stimulus, peaking at around 116 ms for the auditory localizer task (Figure 2.2B, orange curve), 148 ms for the preparatory phase of the attention task (Figure 2.2B, purple curve), and 112 ms when training the classifier on the localizer task and testing on the attention task (Figure 2.2B, pink curve). In both sensor space (Figure 2.2B) and all ROIs (Figure 2.2C-D), cue decoding during the attention task returned to chance level. During the auditory localizer task, cue decoding was more sustained, especially in the LPFC. After matching the number of trials used to train the classifier, an analysis type × ROI ANOVA of peak decoding accuracies, within a 0 – 600 ms time window, showed a main effect of ROI ($F(2,34) = 155.2$, $p < 0.001$), but no differences in decoding amplitude ($F(2,34) = 3.4$, $p > 0.5$), and no interactions ($F(4,68) = 0.7$, $p = 0.6$). Therefore, we found no evidence for template representation beyond the initial auditory representation of the cue.

To test whether activity during any stage of the preparatory phase might reflect the representation of the upcoming trial target, we performed a cross-task and cross-time classification analysis trained using the visual localizer task. At every time window, patterns from the two visual items associated with each cue were taken from the visual localizer task to

use as training data, and these were tested at every time window of the preparatory phase of the attention task to decode the trial target (now without subsampling trials). We did not find any significant time points where the visual template cross-generalized to the preparatory phase.

Finally, we note that cross-time generalization matrices suggest that the LPFC signal reached a steady state at the end of the auditory localizer, in contrast to its lack of any sustained signal during the preparatory phase of the attention task. Even including all the trials of the attention task, without subsampling, we observed the same disappearance of cue decoding during the preparatory phase (see Appendix A. Figure A.7). This might reflect the fact that the representation in the auditory task does not need to be transformed further, whereas in the attention task it serves an intermediate role in mapping subsequent visual inputs to behavior.

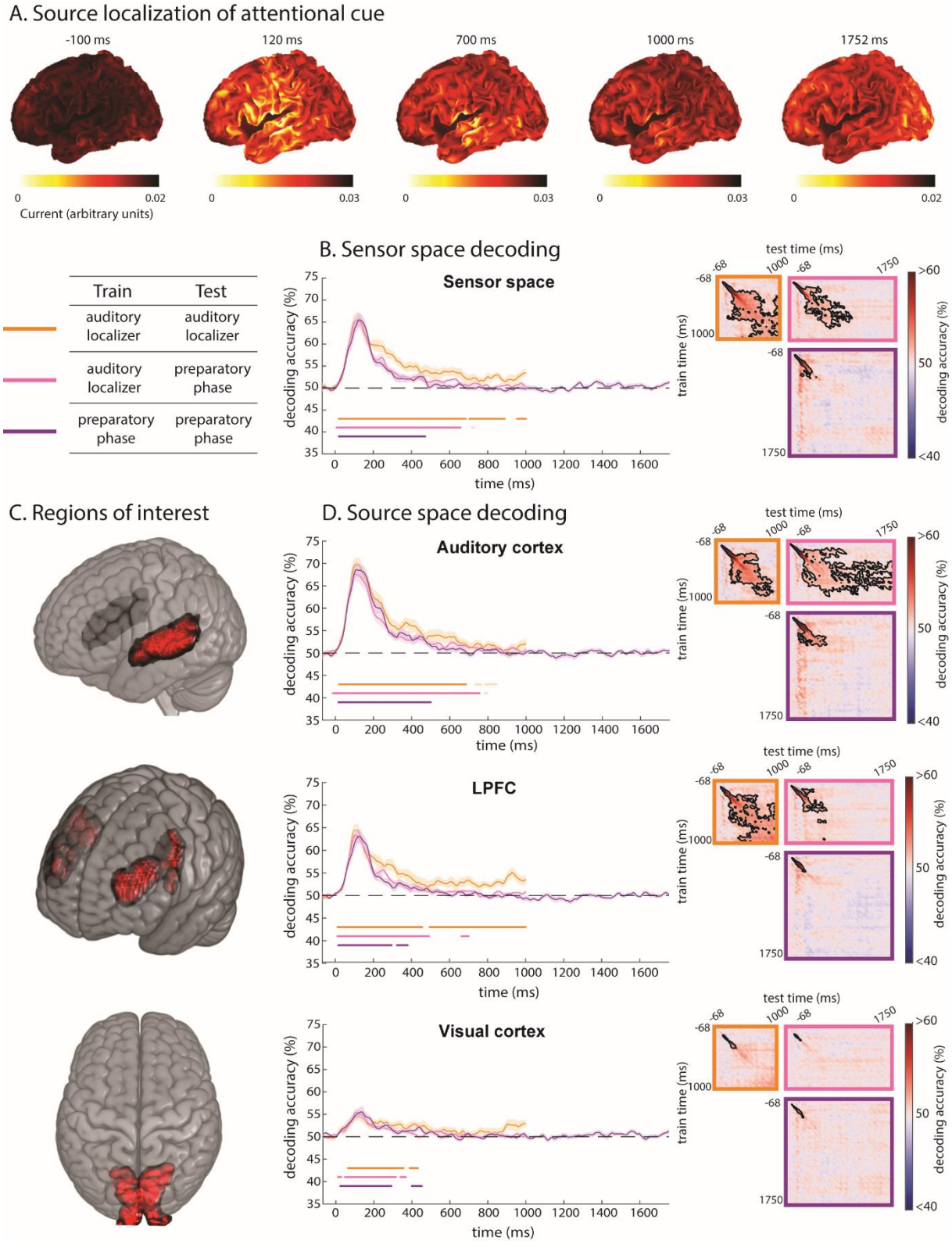


Figure 2.2. Response to the attentional cue. (A) Source localization of EEG/MEG response to the auditory attentional cue at representative time points relative to cue onset. (B) Decoding time-course of auditory stimulus/attentional cue using all sensors combining EEG and MEG across the whole brain. Curves on the left show decoding when training and testing on matched time-points. Dark colored dots beneath the decoding curves show times where decoding is significantly above chance for each condition

($p < 0.05$), corrected for multiple comparisons along the diagonal of the cross-temporal generalization matrix; faint colored dots represent additional time-points where the diagonal of the cross-temporal generalization matrix is significant when corrected for multiple comparisons across the whole matrix. Translucent bands represent standard error of the mean. Matrices on the right show temporal generalization of decoding across all pairs of training and testing times. Black contours indicate regions of significant decoding ($p<0.05$). (C) Vertices within source space ROIs (auditory cortex, lateral prefrontal cortex (LPFC), and visual cortex). (D) Decoding time-courses from these source space ROIs; same format as (B). Significance is corrected for multiple comparisons across time using TFCE and permutation testing.

2.4.3 Coding of visual and behavioral properties of 1-item displays

We next turned to processing of the visual items, and selection of the target item. Source localization of the response to the visual stimuli at representative time points is shown in Figure 2.3A.

During 1-item displays, we expected strong, early discrimination of object identity (e.g., face vs. house, when the consistent non-target was the violin). In the attention task, each stimulus additionally had a behavioral category depending on the cue of that trial. For the participant to make the appropriate response to each stimulus, we expected that the neural signal would also show behavioral category discrimination (target vs. non-target), which would occur after object identity processing. For these analyses, we focused on the T and Ni conditions, for which object identity and behavioral category were fully crossed. Object representation was measured by the discrimination between stimulus identities (e.g. face vs. house) when each were equally often targets or non-targets; conversely, behavioral category representation was measured by discrimination between targets and non-targets when these were equally balanced across stimulus identities.

Single stimulus decoding time-courses on T and Ni presentations are shown in Figure 2.3B-C. In line with expectations, both object identity and behavioral category showed substantial periods of significant decoding accuracy. Across the whole sensor space, a significant difference between object identities peaked at around 128 ms. Behavioral category decoding emerged later, slowly rising to a peak at 360 ms.

Source space analysis showed that both types of information could be decoded from all three ROIs. Decoding of object identity in the auditory ROI warns of possible signal leakage

between regions. Visual cortex, however, had the highest decoding accuracy for object identity, while ROIs did not statistically differ in their strength of decoding accuracy for behavioral category.

Cross-temporal generalization indicated that object identity representation was most stable in the visual ROI. In contrast, behavioral category representation was most stable in the LPFC ROI.

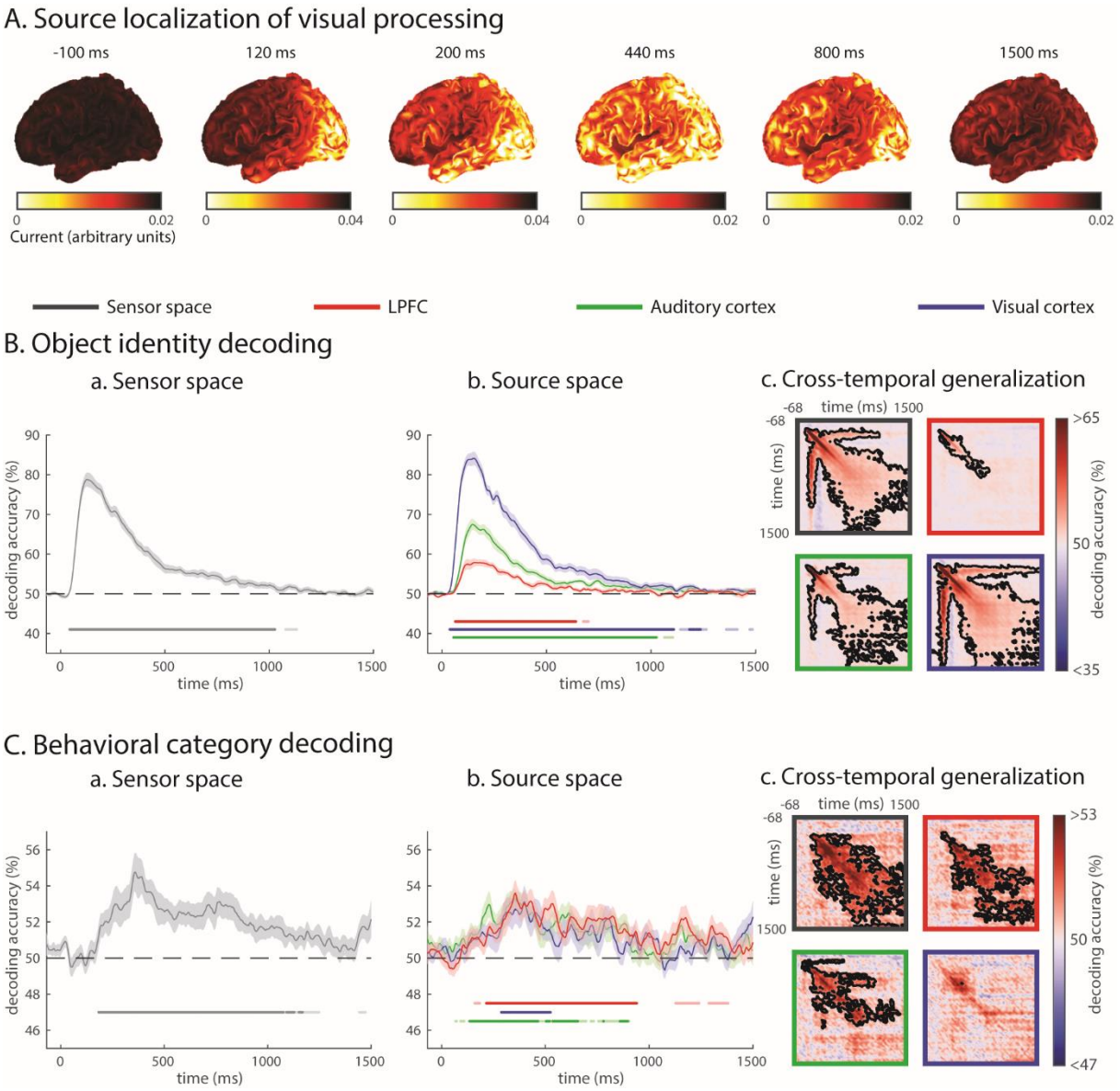


Figure 2.3. Coding of visual and behavioral properties of 1-item displays (A) Source localization of EEG/MEG response to visual presentation (including both single-item and multi-item displays) at representative time-points. (B) Decoding time-courses of object identity, in (a) sensor and (b) source space, when training/testing using matched time-points, and (c) generalizing across training/testing

times. Dark colored dots beneath the decoding curves show times where decoding is significantly above chance for each condition ($p < 0.05$), corrected for multiple comparisons along the diagonal of the cross-temporal generalization matrix; faint colored dots represent additional time-points where the diagonal of the cross-temporal generalization matrix is significant when corrected for multiple comparisons across the whole matrix. Translucent bands represent standard error of the mean. Black contours in cross-time matrices indicate regions of significant decoding ($p<0.05$). Significance is corrected for multiple comparisons across time using TFCE and permutation testing. (C) Decoding time-courses and cross-temporal generalization for behavioral category information. Object identity decoding emerged earlier than behavioral category decoding. Visual cortex showed the highest object decoding accuracy, while ROIs were comparable in their strength of behavioral category representation.

2.4.4 Coding of target location in 3-item displays

Next, we examined target representation in the presence of simultaneous distractors. We first asked when the spatial location of the target within 3-item displays could be decoded (Figure 2.4; Fahrenfort et al. 2017). To do this, we decoded every pair of T versus Ni locations, while holding Nc position constant (i.e., "T right, Ni left" vs. "Ni right, T left", "T right, Ni bottom" vs. "Ni right, T bottom", and "T left, Ni bottom" vs. "Ni left, T bottom") and averaged the accuracies within each participant. Within each pair, collapsing across both possible cues ensured that the decoding was balanced for both visual features and auditory cues. Group sensor-space results showed that decoding began to emerge shortly after stimulus onset, and peaked at 244 ms, before slowly declining toward the end of the epoch. The analysis was repeated in source space. Decoding of target location was significant in all ROIs, but strongest in visual cortex where it peaked at 132 ms. Cross-temporal generalization suggested that the representation of target location was initially dynamic, then entered a temporarily stable state, most apparent in sensor space suggesting spatially coarse stability, before becoming unstable once more prior to the end of the epoch.

In a complementary analysis to target location decoding, we examined the N2pc, a well-known early index of spatial attention, which appears as a negativity over posterior EEG electrodes contralateral to the side of space to which the subject is attending around 200-300 ms following a stimulus (Luck and Hillyard, 1994; Heinze et al., 1990; Eimer, 1996; Hopf et al. 2000; Fahrenfort et al., 2017). We compared event-related potentials/fields when the target was on the right or left of the screen of the 3-item display, and the topography of this contrast

is shown in Figure 2.4B. Differences between target locations peaked between 200-300 ms in posterior EEG and MEG signals, although the signals diverged earlier in MEG, which could reflect the source of the earlier decoding. We note that our lateralized stimuli were in the upper visual field, and that the N2pc is typically stronger for stimuli in the lower visual field (Luck et al 1997; Bacigalupo & Luck, 2018).

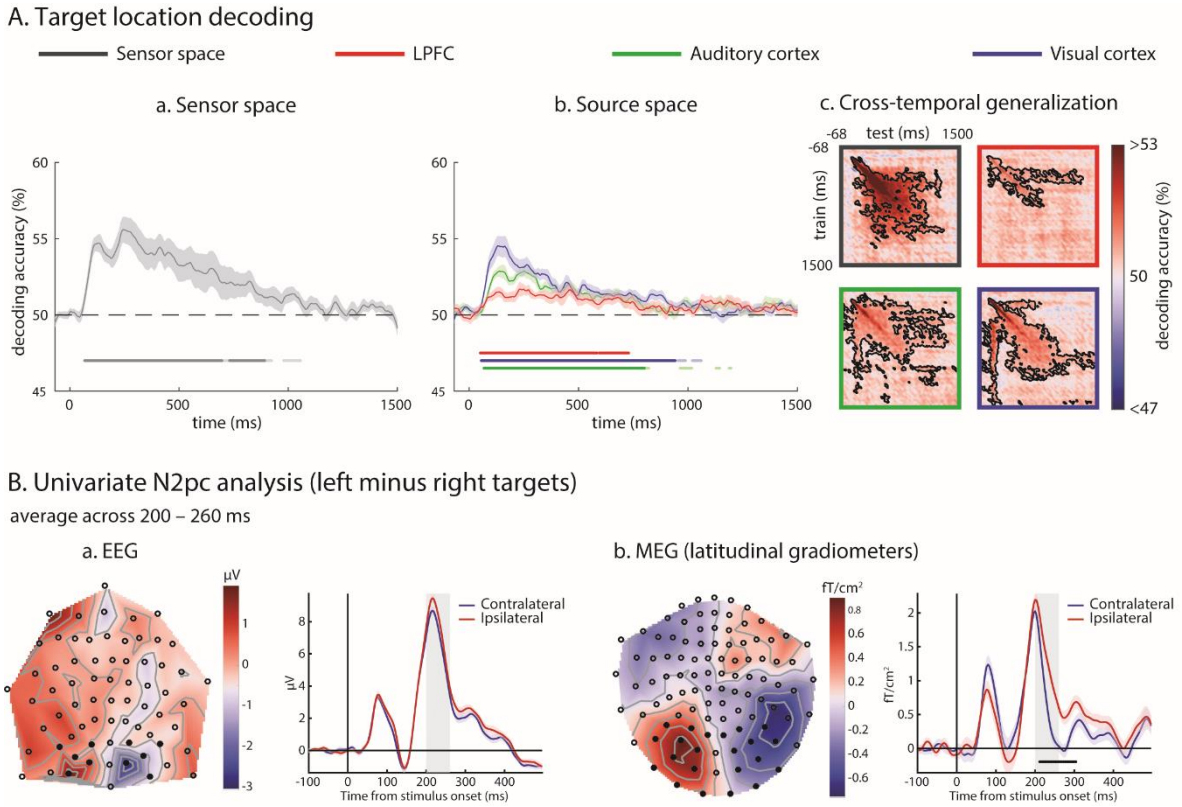


Figure 2.4. Coding of target location in 3-item displays. (A) Decoding of target location during presentation of 3-item displays, i.e., whether the item corresponding to the cue is in the left, right, or bottom position. Format as in Figure 2.3. Location decoding was strongest in the visual cortex. (B) Univariate N2pc ERP/ERFs across (a) EEG electrodes and (b) latitudinal gradiometers. Latitudinal gradiometers are presented because their orientation around the helmet means that contralateral asymmetries in the magnetic flux gradient are expressed analogously to the EEG topography (Mitchell and Cusack, 2011; Kuo et al. 2016). Topographies are averaged across 200-260 ms (marked in grey on the time-courses). Time-courses are averaged across posterior sensors contralateral and ipsilateral to the target (highlighted on the topographies), with black dots indicating a significant difference ($p<0.05$) after TFCE with permutation testing.

2.4.5 Coding of target identity during presentation of 3-item displays

We also hypothesized that representation of 3-item displays would differ depending on the cue, even though the visual input was the same. All 3-item displays contained the target item that was associated with the cue, as well as the Ni and Nc items. Therefore, the decoding of the cue in the presence of a matching visual stimulus likely reflects attentional enhancement of the selected target identity. Although a template representation could also contribute to the decoding, this can only be isolated in the absence of a target (see next section). In sensor space, cue/target identity decoding peaked at 252 ms. In source space, the visual cortex showed the highest decoding accuracy (Figure 2.5).

Cross-temporal generalization suggested that the representation of target identity in the presence of distractors was dynamic, and decayed rather than settling into a steady state. For this analysis, we also expected cross-task generalization from the visual localizer. This was significant in the visual ROI, but not in the auditory or LPFC ROIs, suggesting that the shared pattern was predominantly sensory, with minimal signal leakage in this case.

To compare the decoding latencies of target location and target identity in 3-item displays, we calculated 50%-area latency (Luck, 2014; Liesefeld, 2018) using data from a 0 – 600 ms window for each subject, ROI and decoding type. Paired 2-tailed t-tests showed that target location decoding preceded target identity decoding in both the whole sensor space ($t(17) = 2.86$, $p < 0.05$) and in the visual cortex ($t(17) = 4.97$, $p < 0.001$), but not in the auditory or LPFC ROIs (both $t(17)<1.95$; both $p>0.05$).

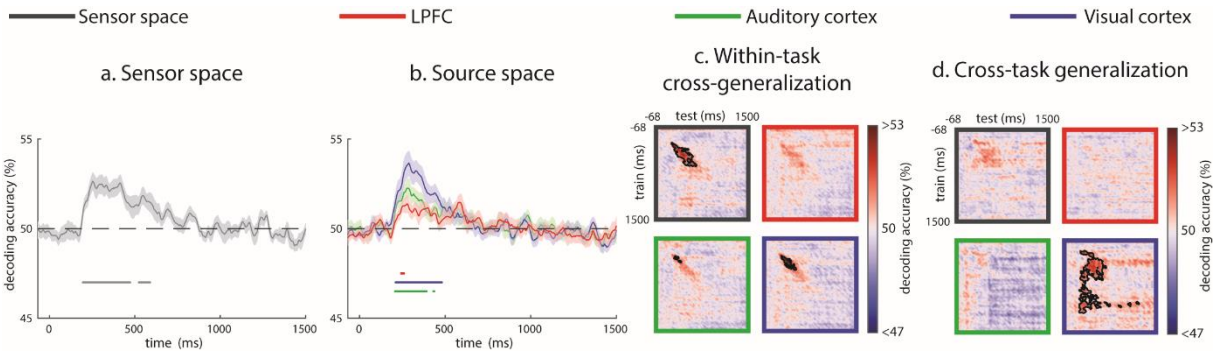


Figure 2.5. Decoding of attentional cue/target identity during presentation of 3-item displays. Panels (a-c) have the same format as Figure 2.3. Panel (d) shows cross-task generalization of decoding, when training on the visual localizer task and testing on the attention task.

2.4.6 Reawakening of the attentional cue/template during presentation of consistent non-targets

Finally, we tested whether we could decode the cue/template during the presentation of a single Nc visual stimulus. Wolff et al. (2015, 2017) have shown that by 'pinging' the brain with a neutral stimulus during working memory maintenance, it can be possible to decode the memory-item-specific information from the impulse response. In our data, cue decoding following Nc presentation was visible but rather weak and intermittent (Figure 2.6). Across sensor space and source space, there were scattered brief periods of above-chance decoding. Their appearance in auditory as well as visual and frontal ROIs questions whether these might reflect a reactivated memory of the auditory cue, or a visual attentional template in anticipation of the next visual input. Apparent signal in the auditory ROI might also reflect leakage from other sources. Cross-temporal generalization suggested that although the representation was not fully sustained, when it resurfaced in the visual ROI it did so with a similar pattern. Cross-task generalization from the auditory and visual localizers provided no evidence that this representation was in a similar format to either cue or target perception.

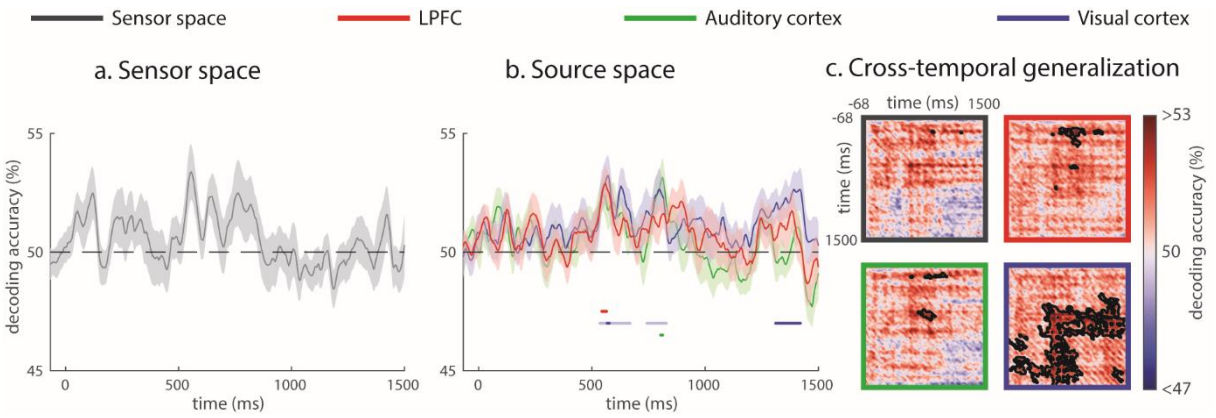


Figure 2.6. Decoding time-course of attentional cue during presentation of Nc displays. Format as in Figure 2.3.

2.4.7 Summary of component time-courses during attentional selection

Above we have described five distinct forms of information representation evoked by the appearance of the visual stimuli (Figures 2.3-6). These are summarized in Figure 2.7, overlaying their average sensor-space and ROI-based decoding time-courses for ease of comparison.

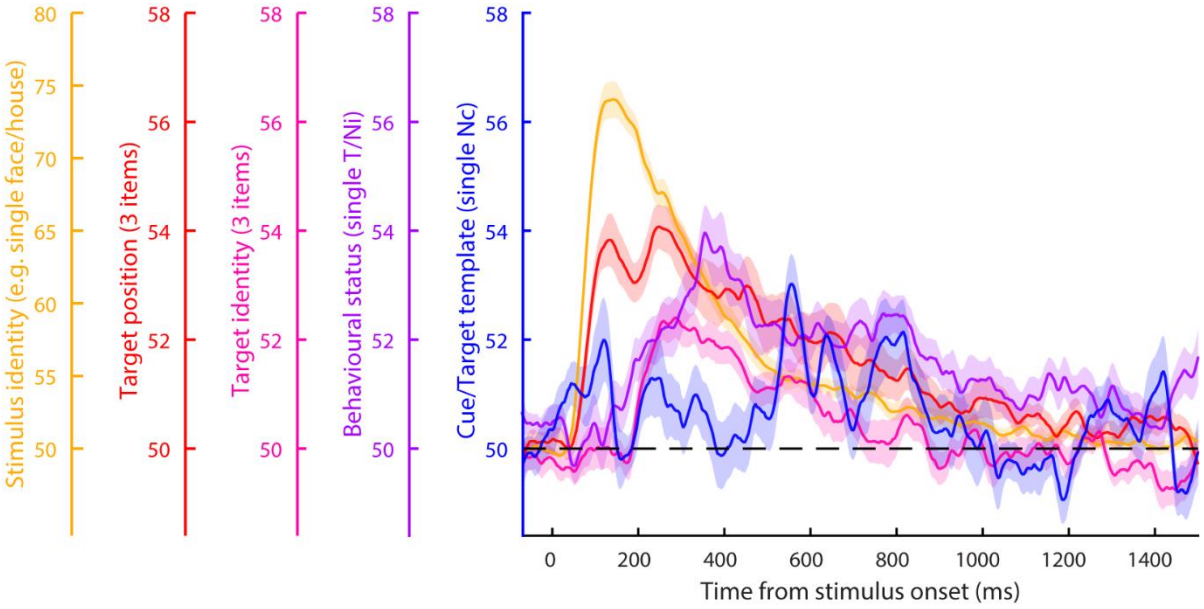


Figure 2.7. Summary of the decoding time-courses of five component processes of selective attention following onset of a visual stimulus: representation of stimulus identity, target position, target identity, behavioral status, and the template of the cue/target. Decoding accuracy is averaged across sensor space and source ROIs, and translucent bands represent standard error of the mean across subjects.

2.5 Discussion

There is currently much interest in decoding the contents of cognitive operations from human MEG/EEG data, and in using these methods to understand attentional selection of information relevant to current goals. Here, we examined the evolution of multiple forms of information represented in the brain as a visual target is selected. Combining single-item displays with multi-item displays of targets and different types of distractors allowed quantification of distinct components of processing during selective attention, indexed by different profiles of representational content.

Although multiple attentional templates could guide behavior (Awh et al., 2012), for effective task performance selection of a particular target requires a template that specifies the currently relevant object (Duncan and Humphreys, 1989; Bundesen, 1990). In fMRI, multivariate classifiers trained on responses to viewed stimuli can predict an attentional template during the preparatory phase (e.g., Stokes et al., 2009). In our MEG/EEG data, we observed significant decoding of cue identity in the attention task, but after equating trial numbers decoding accuracies were not significantly different from that of stimulus processing

in the auditory localizer task. Furthermore, beyond 1000 ms, cue decoding was indistinguishable from chance. Previous MEG/EEG studies have suggested existence of a pre-stimulus template, often subtle and short-lived (Myers et al., 2015; Kok et al., 2017; Grubert and Eimer, 2018). Following non-target displays, we observed evidence of template reawakening; although significant, this was weak and not fully sustained. The delay of template reactivation relative to the explicit categorization of the display as a non-target suggests a serial component to the search process, here within the temporal presentation stream but consistent with neural evidence of serial refocusing of attention within single search displays (Woodman & Luck, 1999; Bichot et al., 2005).

Sustained preparatory activity reflecting an attentional template may be largely invisible to MEG/EEG for many reasons. For example, at the physiological level, if discriminating neurons are intermixed, they may be hard to distinguish with non-invasive methods. Recent findings from single trial analysis of direct neural recordings also suggest that spiking activity during the delay period is sparse, with brief bursts of activity having variable onset latency and duration, which would hinder cross-trial decoding (Shafi et al., 2007; Lundqvist et al., 2016a, 2018; Stokes and Spaak, 2016; Miller et al., 2018). A parallel possibility is that attentional templates may sometimes be stored in an "activity silent" passive form, such as changed synaptic weights (Lundqvist et al., 2010; Stokes, 2015). Consistent template representations may also be difficult to detect if there is-trial-to-trial variability at the cognitive level (Vidaurre et al., 2019), such as fidelity of mental imagery, as well as the anticipation of stimulus timing, with templates activated/strengthened only when the search display is expected to be imminent (Grubert and Eimer, 2018). It is also possible that in the current experiment, the attentional template required little effort to maintain as a verbal label and might have been more visible if harder to verbalize. Consistent templates may be more likely when few features distinguish targets from distractors, for example when targets are defined only by orientation or color (Kok et al., 2017; Myers et al., 2015; Grubert and Eimer, 2018). Perceptually complete templates may be more likely when targets share different features with different distractors (Duncan and Humphreys, 1989). Finally, we emphasize that for successful task performance a template must exist in some form, even when we are unable to detect it, and that uncovering subtle or variable templates may benefit from novel analysis methods (Vidaurre et al. 2019).

Upon presentation of the visual choice display, we found much decodable information of various kinds. The timing of peak decoding of different features suggests five components of processing. The current data cannot determine the extent to which these components evolve

in parallel or have some serial dependency, whereby one process influences another. It is likely that there is a degree of both (Bichot et al., 2005). First, visual stimulus properties are encoded, shown by object identity decoding in 1-item displays, peaking around 128 ms, and strongest in visual cortex. Second, in a multi-element display, the candidate target is localized, shown by target location decoding that peaked between 132 ms (in visual cortex, where strongest) and 244 ms (combining all sensors). This may be partially concurrent with initial visual processing, consistent with an initial parallel stage of selection (Duncan, 1980; Treisman and Gelade, 1980) and automatic registration of coarse feature location (e.g. Cohen & Ivry,1989; Hopf et al., 2004), that could be used to guide subsequent attention (Itti and Koch, 2000; Bisley and Goldberg, 2010; Wolfe, 1994; Eimer, 2015). Third, representation of the candidate target continues to be enhanced relative to distractors, perhaps via integrated competition, shown by cue/target identity decoding in 3-item displays, peaking around 252 ms, again strongest in visual cortex. Fourth, behavioral significance of the target is explicitly represented (in this case whether it is a target, so requiring further processing), shown by behavioral category decoding in 1-item displays, peaking around 360 ms and most stable in the LPFC. Fifth, if no target is identified and search must continue, an attentional template might be reactivated or strengthened, shown by cue decoding after Nc displays, peaking beyond 500 ms. The precise timing at which each representation is detectable will depend on many factors including stimuli, task, analysis sensitivity, similarity between targets and distractors, and the number and homogeneity of distractors (Duncan and Humphreys 1989). Nonetheless, we anticipate that the sequence of key components would largely generalize across paradigms (Eimer, 2015; Vidaurre et al., 2019). Potential dependencies between processes might be investigated by combining MVPA of electrophysiological recordings with transcranial magnetic stimulation at successive times.

In 1-item displays, we found a distinction between visual cortex and LPFC. While the regions represented behavioral category with similar strength, visual cortex represented stimulus identity more strongly than LPFC. Similarly, object identity was represented more stably in visual cortex, whereas behavioral category was represented more stably in LPFC. fMRI studies show that frontal regions flexibly code for behaviorally relevant categories according to task rule (Jiang et al., 2007; Li et al., 2007; Woolgar et al., 2011a; Lee et al., 2013; Erez and Duncan, 2015). Electrophysiological recordings of monkey prefrontal responses to T, Ni, and Nc stimuli show that visual input properties are initially equally represented for targets and non-targets, whereas the behaviorally critical target dominates later processing (Kadohisa

et al., 2013; Stokes et al., 2013). Our results also suggest an anterior-posterior distinction in information content and timing.

In multi-item displays, the candidate target was rapidly identified and localized, with location decoding providing the earliest evidence of modulation by behavioral relevance. Although its timing, peaking around 132 ms in V1, was earlier than might be expected based on the N2pc and multivariate decoding using EEG alone (Fahrenfort et al., 2017), it is consistent with representation of the location of task-relevant features reported from ~140 ms and preceding the N2pc (Hopf et al. 2004). Ipsilateral and contralateral target responses diverged earlier in MEG than EEG, suggesting that the source of the earlier decoding may be more visible to MEG. Location decoding peaked later in the other ROIs and at the sensor level (beyond 230 ms) suggesting that source localization may have helped in isolating the earlier signal.

Although target localization implies target identification, and time-courses of location and identity representation in 3-item displays were heavily overlapping, the location signal was significantly earlier than the identity representation in visual cortex. This is consistent with models of visual attention as well as empirical data that make an explicit distinction between feature selection, where attention is rapidly allocated to candidate objects (Broadbent, 1958), and object recognition, which takes place at a subsequent stage where the features of objects are integrated and their identity becomes accessible (Eimer, 2015; Eimer & Grubert, 2014; Kiss et al., 2013). It could also arise within a continuous competitive framework, without explicit recognition, if neurons representing identity have overlapping receptive fields such that competition amongst them is slower to resolve or benefits from prior spatial filtering (Luck et al. 1997); or if complete identity representation involves several features whose integration is strongly mediated by shared location within spatiotopic maps (Treisman & Zhang, 2006; Schneegans and Bays, 2017). The location of an attended feature can also be represented before the location of a target itself (Hopf et al., 2004), and the temporal priority with which different features of the target are enhanced may depend on the cortical location as well as the particular task demands (Hopf et al., 2005). The observations that competitive representations of target location and target identity peaked at different times, and that neither appeared to reach a permanent steady state, together indicate that the early phase of integrated competition is dynamic, with different aspects of the target representation waxing and waning at different times. In contrast, the later explicit representation of target status settled into a steady state in LPFC that persisted until the end of the epoch.

Interestingly, a target influenced bias in the 3-item displays well before its target status was explicitly decodable in the single-item displays. This strongly suggests at least two stages of target processing, consistent with behavioral manipulations suggesting that spatial selection and target identification are separable (Ghorashi et al., 2010). Distinction between an early, parallel processing stage and a later capacity-limited stage is central to most models of attention (Duncan, 1980; Treisman & Gelade, 1980). Target decoding in 3-item displays peaked at 252 ms with first significance at 196 ms, similar to attentional modulation of stimulus category processing in cluttered scenes observed from 180 ms (Kaiser et al., 2016), and to demonstration of feature-binding during integrated competition (Schoenfeld et al. 2003). The later stage indexed by single-item decoding may correspond to capacity-limited individuation of the integrated target object, allowing its bound properties to become accessible for further processing and goal-directed action (Duncan, 1980; Bichot et al., 2005; Mitchell and Cusack 2008; Christie et al., 2015), in this case likely including the brightness judgement. These two stages could also be interpreted in terms of the "global neuronal workspace" model - the earlier attentional bias reflecting accumulation of pre-conscious sensory evidence; the later explicit representation of target status reflecting conscious awareness and "ignition" of fronto-parietal networks, linked to P3 waves around 300- 500 ms (Dehaene & Changuex, 2011; Sergent et al., 2005) and consistent with the timing of peak decoding at 360 ms.

To conclude, although attentional selection must begin with a template, this may be weakly or variably represented (Duncan et al., 1997; Lundqvist et al., 2018; Miller et al., 2018), such that it is largely invisible to MEG/EEG, or even maintained in "silent" form (Stokes, 2015). In agreement with others (Olivers et al., 2011; Myers et al., 2015; Grubert and Eimer, 2018), we suggest that the template may be actively and consistently represented only when needed, and least likely to interfere with other concurrent processes. Integrated competition accounts of attention imply that the template need be neither complete nor constant across trials, consistent with no significant response pattern generalization between template representations and the visual localizer. In contrast, integrated competition suggests that attentional selection and enhancement of stimulus representations will be strong and widespread. Supporting such models, we observed robust, time-resolved decoding of the critical processing stages required to select and enhance a target amongst competing distractors, and to categorize it according to behavioral requirements.

Chapter 3 Representation of task episodes in human cortical networks

3.1 Abstract

Task episodes consist of sequences of steps that are performed to achieve a goal. The current study used fMRI to examine which regions of the brain represent full episodes, component items, and sequential position. Participants learned six tasks each consisting of four steps. Inside the scanner, participants were cued which task to perform and then sequentially identified the target item of each step in the correct order. The multiple demand (MD) network and the visual cortex exhibited phasic responses to each task step, suggesting that they are sensitive to the fine structure of the episode. In addition, many brain regions – including regions of the default mode network (DMN) - showed a phasic response to onset and offset of the entire task episode, along with gradually increasing activity across the episode. Representational similarity analysis was used to examine encoding of component items and entire episodes. Compared to MD regions, which coded individual items but not the entire episode, the DMN showed representation of both item and episode. The results suggest collaboration of multiple brain regions in control of multi-step behavior, with MD regions representing the detail of individual steps, and DMN adding representation of broad task context.

3.2 Introduction

A central feature of purposeful everyday behavior is the retrieval of learned sequences of events from memory (Hsieh and Ranganath, 2015) to guide our current actions. This involves parcellating a main goal (e.g., “make a stew”) into smaller achievable steps (e.g., “take food from fridge” → “wash vegetables” → “chop vegetables” → “cook on stove”) to allow progression towards the goal (Penfield and Evans, 1935; Cooper and Shallice, 2000; Farooqui et al., 2012). We call these temporally organized sequences of steps that occur within a given context “task episodes”. A key aspect of these task episodes is the control of extended episodes of behavior as one unit, and not as a collection of independent acts (Schneider and Logan, 2006; Duncan, 2010; Farooqui and Manly, 2018b). Whenever a step is completed, its specific content loses relevance, but higher level task representations of the full episode must remain in

behavioral control (Farooqui et al., 2012; Farooqui and Manly, 2018b). This raises the question of how different brain regions work together to execute the current step of the task while keeping the overall goal in mind.

Previous literature has highlighted the importance of a set of frontal and parietal regions, known as the multiple demand (MD) network (Duncan and Owen, 2000), in executing complex mental programs (Duncan, 2010, 2013; Farooqui et al., 2012). It has been proposed that the MD network plays a key role in defining and controlling parts of task episodes, allowing goals to be achieved by decomposition into a structure of subgoals (Kurby and Zacks, 2008; Farooqui et al., 2012; Duncan, 2013). The MD network is well suited for focusing on specific contents of a current cognitive operation, dynamically encoding information relevant to a current decision (Asaad et al., 2000; Everling et al., 2002; Li et al., 2007; Woolgar et al., 2011b; Stokes et al., 2013), and radically changing the pattern of activity across successive task steps (Sigala et al., 2008; Duncan, 2010). In particular, Farooqui et al. (2012) investigated the role of MD activity in task episodes requiring a series of target detection steps. The authors found that target detections that completed the entire task episode elicited the greatest MD activity, followed by those completing a subtask, and finally steps within one subtask. As MD activity depended on task completion, it was suggested to be involved in directing and revising the control representations of each step of the episode.

The ability to organize sequences of events within a given context has also been a key topic in the study of episodic memory (Ezzyat and Davachi, 2011; Eichenbaum, 2013; Hsieh et al., 2014; Cohn-Sheehy and Ranganath, 2017; Radvansky and Zacks, 2017). Tulving's original definition emphasized the importance of temporal events: "Episodic memory receives and stores information about temporally dated episodes or events, and temporal-spatial relations among these events (Tulving, 1972, p.385)." Event segmentation theory (Zacks and Tversky, 2001; Zacks and Swallow, 2007; Radvansky and Zacks, 2017) proposes that humans can segment incoming information into temporal parts that are meaningfully related to the current situation. When important situation features change, the current event model is updated and experienced as an event boundary. Neuroimaging studies have found brain regions sensitive to event boundaries to overlap with areas associated with episodic memory retrieval including regions in the default mode network (DMN; Zacks et al. 2001; Speer et al. 2007; Ben-Yakov et al. 2014; Richmond and Zacks 2017; Baldassano et al. 2018). Furthermore, it has been suggested that these dynamics within the DMN may reflect the underlying meaning of the episode rather than simple stimulus changes (Radvansky and Zacks, 2017), as coarse

segmentation elicited greater DMN activity than fine-grained segmentation (Speer et al., 2007). Consistent with this observation, the DMN network has been implicated in high level cognition at a broad scale, such as encoding of schemas (Robin and Moscovitch, 2017), situation models (Reagh and Ranganath, 2018), and cognitive contexts (Crittenden et al., 2015). In a study of topographic mapping of a hierarchy of temporal receptive windows (TRW), participants listened to a story scrambled at the time scales of words, sentences, and paragraphs (Lerner et al., 2011). Results showed that early sensory regions were driven by incoming sensory input and were similarly responsive in all conditions; however, MD regions exhibited intermediate TRWs, whereas DMN regions were at the apex of the TRW hierarchy, such that they responded reliably only when intact paragraphs were heard in a meaningful sequence. This evidence suggests that the DMN is well suited to representing task episodes over an extended timescale.

Although various brain networks have been implicated in the execution of task episodes, to our knowledge, no study has contrasted the roles of MD and DMN in representing different aspects of a task episode. In the current study, we aimed to examine which brain regions are involved in coding of information at various levels of abstraction within a single task: individual steps, including their content and position within an episode, whole episodes, and groups of related episodes. Prior to the experiment, participants learned six everyday task episodes associated with different rooms (three kitchen tasks and three bathroom tasks) that each consisted of four steps. Inside the scanner, after being cued which task to perform, participants sequentially identified the target item of each step in the correct order. This design allowed us to examine which brain regions represent rooms (e.g., kitchen), full episodes (e.g., "make a stew"), items within the episode ("take food from fridge"), and current position in the episode (e.g., first step). We hypothesized that different regions of the brain would be sensitive to different levels of the temporal task hierarchy. We first focused on the MD and DMN networks as *a priori* regions of interest. We hypothesized that the MD network would be especially involved in moment-to-moment control (Duncan, 2010, 2013; Farooqui et al., 2012); whereas the DMN would be especially involved in representation of full episodes (Lerner et al., 2011; Reagh and Ranganath, 2018). In addition to these pre-defined networks, we examined information coding using a whole brain searchlight to localize relevant brain regions at a finer scale, both within and beyond the *a priori* networks. The visual cortex encodes physical properties of visual stimuli, and therefore we expected it to be involved in item coding along with the MD system. It has been suggested that both the MD network (Dosenbach et al., 2006, 2007) and the DMN (Andrews-Hanna, 2012; Ranganath and Ritchey, 2012) can be divided into

finer components or subsystems, therefore we additionally examined subregions within each network separately. We used both univariate finite impulse response (FIR) models to characterize the temporal evolution of activity though the extended episode, and representational similarity analysis (RSA) to investigate coding of cognitive representations of task structure and content. This was done on large-scale networks, their component ROIs, and at the whole brain level with a searchlight approach in the RSA analysis.

3.3 Methods

3.3.1 Participants

42 participants (20 male, 22 female; ages 18-39, mean = 26.79, SD = 4.77) were included in the experiment at the MRC Cognition and Brain Sciences Unit. An additional 18 participants were excluded (two were discovered to have cysts, one lost several slices due to poor bounding box positioning, ten were excluded due to having no correct episodes for at least one combination of cued task × distractor task, and a further six were excluded due to excessive head motion > 5 mm). All participants were neurologically healthy, right-handed, with normal or corrected-to-normal vision. Procedures were carried out in accordance with ethical approval obtained from the Cambridge Psychology Research Ethics Committee, and participants provided written, informed consent before the start of the experiment.

3.3.2 Stimuli and task procedures

The study consisted of a learning session outside the scanner and an execution session in the scanner. During the learning session, participants learned six everyday task sequences ("episodes") each based in one of two locations ("rooms"; three kitchen and three bathroom). Each episode consisted of four ordered "steps". For example, the episode "make a stew" consisted of the steps "take food from fridge", "wash vegetables", "chop vegetables", "cook on stove". Each step was associated with a unique image ("item"). The complete set of stimuli is shown in Figure 3.1A.

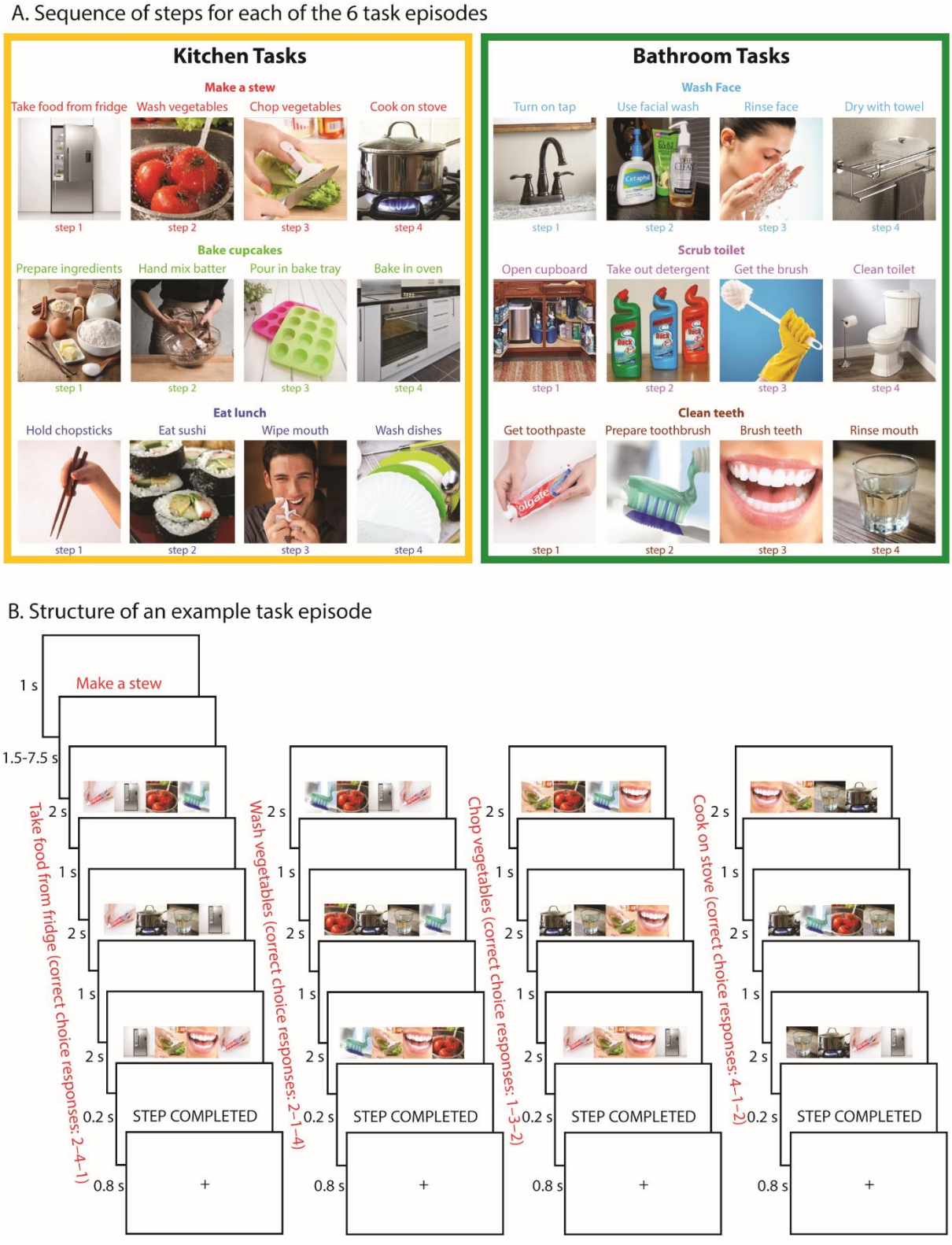


Figure 3.1. (A) Illustration of the six task episodes (three kitchen and three bathroom tasks) memorized before going into the scanner. Each task episode consisted of four steps to be completed in serial order (e.g., the task "make a stew" consisted of "take food from fridge", "wash vegetables", "chop vegetables", "cook on stove"). (B) Structure of an example task episode. Episodes began with a cue indicating which task to perform (e.g., "make a stew").

After a short delay, the first search array of four items appeared, and participants were asked to select the item corresponding to the first step of that task (here, "take food from fridge"). Participants selected this same target in three search arrays (total step duration = 9 s), then were given a brief indicator that the step had been completed, and moved on to the next step (here "wash vegetables"). Completion of all four steps completed the entire task episode.

In the learning session, participants viewed the names and images of the steps of each task episode in sequential order. The step images were presented simultaneously with a background image corresponding to the room in which they occur (kitchen or bathroom). The learning was self-paced, in separate runs for each room. Within each room, each task sequence was presented three times, and each item within the sequence was presented until the participant decided to move on to the next item. There was a 1.5 s inter-stimulus interval between items. After viewing all six sequences, participants were tested for their memory of the task episodes by (1) sorting picture cards representing all steps of the six task episodes into the correct sequences, and (2) completing a pen-and-paper test in which they were asked to write down the names of the steps in the correct order for each task episode. Most participants performed both tests without error. A few participants made a mistake on 1-2 items but were able to correct their answers after being told they make a mistake. The tests ensured participants had memorized the specific step sequence of each task. Before entering the scanner, participants practiced a shortened version of the main experiment, containing one trial of each task episode. During scanning, participants performed two runs of the experiment, interleaved with shorter runs (~5 minutes) of a localizer task that was not analyzed and is not described further.

Figure 3.1B illustrates the structure of the task episodes paradigm. At the start of each 45 s episode, participants were presented with a cue (e.g., "make a stew") for 1 s, indicating which task to complete. This was followed by a fixation period lasting between 1.5 – 7.5 s, selected randomly from a uniform distribution, before the onset of the first step. On each step, participants had to perform three visual searches. On each search, an array of 4 images was presented in a horizontal row (total left to right visual angle approximately 12.6°). These included (randomly ordered from left to right): the correct image ("target") corresponding to the current task step; a distractor image representing a random incorrect step from the correct task; a distractor representing the correct step but from an incorrect task; and an additional distractor from the same incorrect step of the same incorrect task. To ensure that each display contained two images from each room, incorrect-task distractors were selected at random from

the alternative room to the cued task. The array remained for 2 s, and within this time, the participant had to indicate the position of the target image using a 4-choice button box with their right hand. A 1 s fixation interval preceded onset of the next search array. Each step thus lasted for 9 s, with the participant selecting the same target in each of three search events, to allow separation of the hemodynamic response to successive task steps, while ensuring sustained focus on the relevant item within each step. At the end of the third search event, a 0.2 s presentation of the words "STEP COMPLETED" indicated the completion of that step, and was followed by a 0.8 s fixation interval. Without further cueing, the participant then moved on to the next task step. After completing the last step, a fixation interval of 0.5 – 6.5 s was presented before the onset of the cue for the next task. The total interval between the last step of the previous task and the first step of the next task was fixed at 9 s. Participants were not given feedback on their accuracy. Each run consisted of 36 task episodes (with an additional dummy episode to start), constructed so that each task appeared following each possible preceding task once. Task ordering was chosen before the start of each run to maximize the design efficiency (Dale, 1999) of all pairwise contrasts between tasks. 1000 task orders were simulated, and the most efficient one was chosen. Each of the two runs lasted ~28 min.

3.3.3 fMRI data acquisition and preprocessing

Scanning took place in a 3T Siemens Prisma scanner. Functional images were acquired using a multi-band gradient-echo echo-planar imaging (EPI) pulse sequence (TR = 1373 ms, TE = 33.4 ms, flip angle = 74°, 96 × 96 matrices, slice thickness = 2 mm, no gap, voxel size 2 mm × 2 mm × 2 mm, 72 axial slices covering the entire brain, 4 slices acquired at once). The first 5 volumes served as dummy scans and were discarded to avoid T1 equilibrium effects. Field maps were collected at the end of the experiment (TR = 400 ms, TE = 5.19 ms / 7.65 ms, flip angle = 60°, 64 × 64 matrices, slice thickness = 3 mm, 25% gap, resolution 3 mm isotropic, 32 axial slices). High-resolution anatomical T1-weighted images were acquired for each participant using a 3D MPRAGE sequence (192 axial slices, TR = 2250 ms, TI = 900 ms, TE = 2.99 ms, flip angle = 9°, field of view = 256 mm × 240 mm × 160 mm, matrix dimensions = 256 × 240 × 160, 1 mm isotropic resolution).

The data were preprocessed and analyzed using automatic analysis (aa) pipelines and modules (Cusack et al., 2015), which called relevant functions from Statistical Parametric Mapping software (SPM 12, http://www.fil.ion.ucl.ac.uk/spm) implemented in Matlab (The MathWorks, Inc., Natick, MA, USA). EPI images were realigned to correct for head motion using rigid-body transformation, unwarped based on the field maps to correct for voxel

displacement due to magnetic-field inhomogeneity, and slice time corrected. The T1 image was coregistered to the mean EPI, and then coregistered and normalized to the MNI template. The normalization parameters of the T1 image were applied to all functional volumes. The model incorporated a high-pass filter with a cutoff at 1/128 Hz. Spatial smoothing of 10 mm FWHM was applied for univariate analysis, but not for multivariate analysis.

3.3.4 Regions of interest (ROIs)

For the primary analysis, we focused on the MD and DMN networks (see Figure 3.4). The MD network was taken from Fedorenko et al. (2013), and consisted of regions within the lateral prefrontal cortex (LPFC) extending along the anterior, middle, and posterior middle frontal gyrus (aMFG, mMFG, and pMFG), a posterior-dorsal region of lateral frontal cortex (pdLFC), intraparietal sulcus (IPS), anterior insula (AI), and anterior cingulate cortex (ACC). The DMN network was taken from Yeo et al. (2011), combining three subnetworks from the 17 network parcellation (numbers 15, 16, and 17; Andrews-Hanna 2012). The left and right hemispheres were averaged and projected back to both hemispheres to create a symmetrical volume (similar to Fedorenko et al (2013)). The combined networks were then smoothed at 4 mm FWHM, and were split into spatially distinct ROIs, including the medial prefrontal cortex (MPFC) and posterior cingulate cortex (PCC) along the midline, as well as the inferior frontal gyrus (IFG), inferior parietal lobule (IPL), parahippocampal cortex (PHC), and parts of the lateral temporal cortex extending to the temporal pole (Temp). Overlapping voxels of the AI and IFG were excluded from each ROI and their corresponding networks. Analyses were first performed at the network level, and then within each individual ROI to examine more fine scale differences within each network.

3.3.5 Univariate analysis

3.3.5.1 FIR Model

Statistical analyses were performed first at the individual level, using a general linear model (GLM). To capture the BOLD time-course throughout each task episode, as well as transitions between episodes, we modeled each consecutive pair of episodes. The first (dummy) episode was separately modeled and not analyzed. For the remaining data, a 90 s epoch starting from the onset of the first search array of every even number episode to the first search array of the next even number episode was modeled using a finite impulse response (FIR) basis set of 60 1.5 s boxcar regressors. In this way, the response throughout task episodes could be modelled without making assumptions about the shape of the hemodynamic response. Episodes with a

high proportion of errors (episodes that had > 25% errors) were defined as error episodes, with the total number of error episodes per participant ranging from 0-6 (mean = 0.95, SD = 1.43). Any two consecutive error episodes were removed from the analysis using a similar but separate set of regressors. Effects of cues, and errors on individual search arrays, were also modeled, by convolving the duration of their respective events (1 s for cues and 2 s for error events) with a canonical hemodynamic response function. Across the 90 s epoch, estimates for each FIR time bin were extracted from the two networks of interest, averaged over voxels within the network and across the six tasks. These average beta estimates for individual participants were entered into a random effects group analysis.

3.3.5.2 Event-based GLM analysis

To complement the FIR model, an event-based GLM analysis was performed. In this analysis, we aimed to separate phasic activity linked to onset of each step from tonic activity across the whole step within each episode. To control for visual differences, each combination of cued task × distractor task was modeled separately. For each combination, each step was modelled using two regressors, an onset regressor modelled with 0 s duration and an epoch regressor modelled with 9 s duration. Additionally, an offset regressor modelled with 0 s duration was placed at the end of the episode. Each regressor was convolved with the canonical hemodynamic response function. There were accordingly 162 regressors of interest, two (onset and epoch) for each of the four steps and one for each offset of the entire episode in each of combination of six tasks × three possible distractor tasks from the other room (for example, the target episode "make a stew" could be paired with distractor episodes "wash face", "scrub toilet", or "clean teeth"). Error episodes (defined as episodes that had > 25% errors) were removed from the analysis using a similar but separate set of regressors. The cue was modelled using a similar combination of onset (0 s duration) and epoch (duration from cue onset to the onset of the first task step) components. Beta estimates were averaged across the 18 cued task × distractor task combinations for individual participants, and entered into random effects group analyses. We first examined the mean effect of onset/offset and epoch regressors versus implicit baseline (each FDR corrected across number of comparisons). Next, to determine whether the BOLD signal showed significant linear changes towards goal completion, we performed t tests on increasing ([-3 -1 +1 +3]) and decreasing ([+3 +1 -1 -3]) linear contrasts across task steps. To complement the ROI analyses, contrasts were also carried out at the whole brain level, using a voxel-wise FDR-corrected threshold of $p < 0.05$.

3.3.6 RSA analysis

We performed representational similarity analysis (RSA) using the linear discriminant contrast (LDC) to quantify dissimilarities between activation patterns. The analysis used the RSA toolbox (Nili et al., 2014), in conjunction with in-house software. The LDC was chosen because it is multivariate noise-normalized, potentially increasing sensitivity, and is a cross-validated measure which is distributed around zero when the true distance is zero (Nili et al., 2014; Walther et al., 2016). The LDC also allows inference on contrasts of dissimilarities across multiple pairs of task events. A pattern for each step of each combination of cued task and distractor task was obtained, by averaging the onset and epoched responses from the event-based GLM described above. This resulted in 72 patterns in total in each run. For each pair of patterns, the patterns from run 1 were projected onto a Fisher discriminant fitted for run 2, with the difference between the projected patterns providing a cross-validated estimate of a squared Mahalanobis distance. This was repeated projecting run 2 onto run 1, and we took the average as the dissimilarity measure between the two patterns. All pairs of pattern dissimilarities therefore formed a symmetrical representational dissimilarity matrix (RDM) with zeros on the diagonal by definition. To compare dissimilarity magnitude across ROIs of different sizes, the LDC values were normalized by dividing by the number of voxels within each ROI.

3.3.6.1 Coding of information within regions of interest

We first performed this RSA analysis using activation patterns from our *a priori* MD and DMN network ROIs. A simplified version of the resulting 72 × 72 representational dissimilarity matrix (RDM) is shown in Figure 3.2A, with each cell representing a cross-validated LDC dissimilarity between the corresponding two task events. These included event pairs that shared the same episode (red cells; e.g., "take food from fridge" and "wash vegetables"); events that shared the same room but different episodes (purple cells; e.g., "take food from fridge" and "hand mix batter"); and events that differed in both episodes and rooms (orange cells; "take food from fridge" and "use facial wash"). All event pairs additionally differed in item. Saturation of the colors is used to indicate the difference in steps between event pairs. The cells on the diagonal (white) are zero by definition as they do not reflect a distance between different task events. We additionally modelled the amount of visual overlap for each event pair in the RDM as a nuisance regressor (which could be 0%, 50%, or 100%; e.g., an episode with cued task "make a stew" and distractor task "wash face" would have 50% overlap with an episode containing cued task "wash face" and distractor task "bake cupcakes". For the illustration of the full model with the visual confounds, see Appendix B. Figure B.1).

For each ROI of each participant, we fit a linear regression to quantify the influence of differences in room, episode, step, visual difference, and item to the measured pattern dissimilarity (Figure 3.2B). Regressors were created for each possible feature, including episode (0 or 1 indicating same or different episode), room (0 or 1 indicating same or different room), step (step difference 0, 1, 2, or 3), and visual difference (0%, 50%, or 100% different), while the observed LDC values served as the dependent variable. Coefficients were estimated for room, episode, and step, as an index of the strength of representational coding. Representation of visual difference was not of primary interest, but was included as a nuisance regressor to model potential visual confounds. Since all items were unique, item coding could be estimated from the intercept of the model, i.e. the cross-validated pattern dissimilarity in the absence of episode, room, step, or visual differences. Multiple comparisons across ROIs were corrected using FDR < 0.05 for each coding type.

3.3.6.2 Searchlight analyses

Next, to obtain more specific localization of regions that contained information, and to test for additional regions outside the predefined networks, we implemented a whole brain searchlight procedure (Kriegeskorte et al., 2006) to perform pattern analyses in small spherical ROIs (radius = 10 mm) centered on every voxel of the brain in turn. The procedure was identical to that described in the ROI analysis. Pairwise dissimilarities for each cell type were derived from the 72 × 72 RDM in each sphere, and were assigned to the center voxel. This resulted in whole-brain maps of information coding for each subject. These individual subject maps were smoothed with a 10 mm FWHM Gaussian filter before performing second-level random effects analyses to identify voxels that coded for different types of information across subjects. Unless otherwise specified, all results are reported at the FDR-corrected threshold of $p < 0.05$.

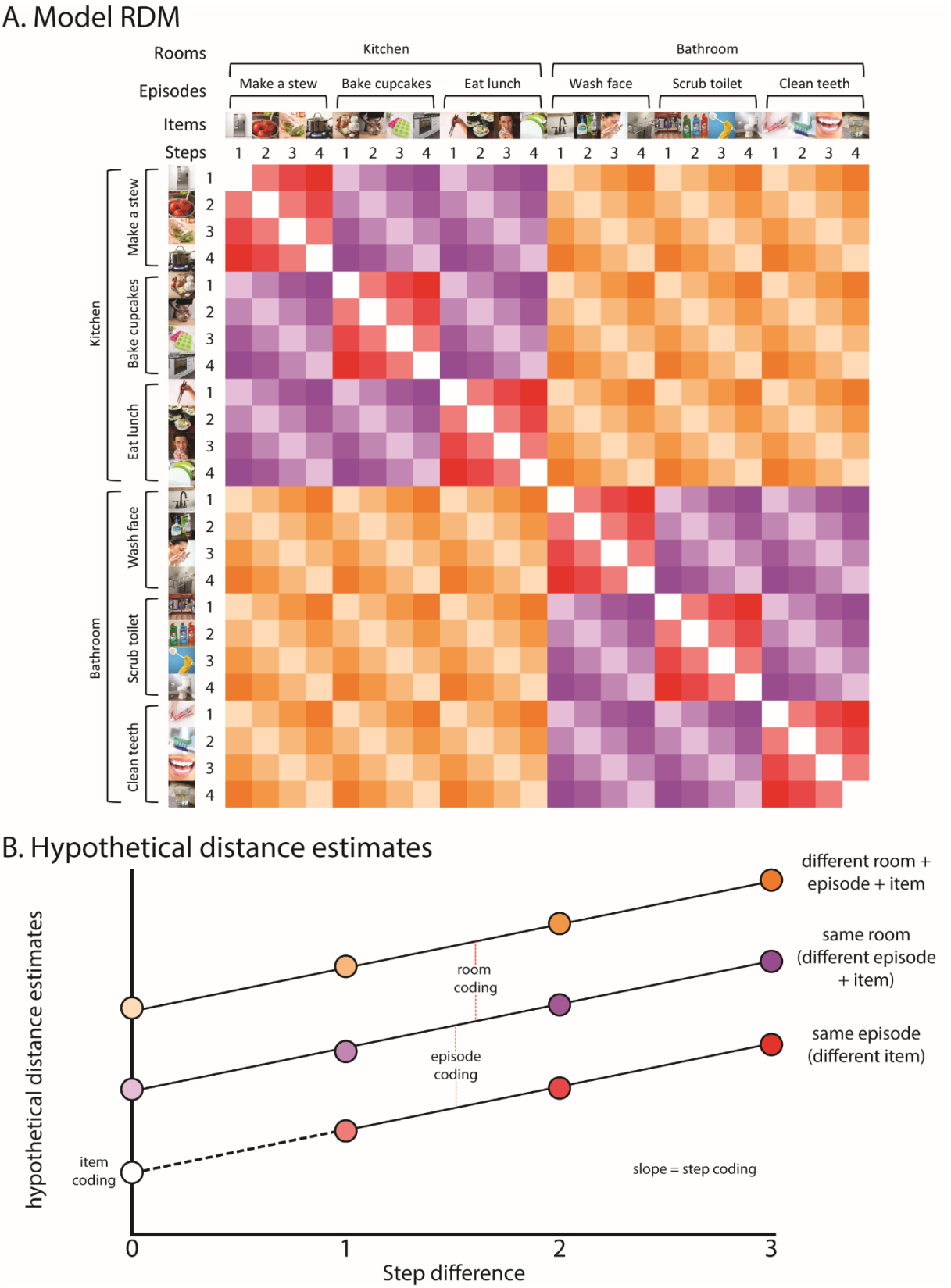


Figure 3.2. Illustration of representational similarity analysis. A. Conceptual RDM. LDC dissimilarities were computed between every possible pair of events (6 cued tasks × 4 steps × 3 distractor combinations per episode), generating a 72 × 72 RDM. Diagonal cells of the RDM

are zero by definition as they do not reflect a dissimilarity between different events. Off-diagonal cells reflect pattern dissimilarity between events that differ in room, episode, step, item, and/or visual difference. B. Hypothetical distances resulting from episode, room, and item coding across step differences. Item coding can be estimated as the intercept, i.e., in the absence of episode, room, step, or visual differences.

3.4 Results

3.4.1 Behavioral results

Group behavioral performance and reaction time results are shown in Figure 3.3. Errors were calculated after removal of error episodes (defined as episodes that had > 25% errors), and reaction time was calculated for only correct trials. Overall accuracy was 97.5% ± 0.4% (mean ± SEM) and overall reaction time was 849 ± 23 ms. Error responses were broken into four error types: choosing an item from same episode different step, different episode same step, different episode different step, and missed response. Results show poorest performance for the first search array of each step, when participants were required to switch from one step to the next. A step (steps 1-4) × search array (first, second, third within each step) ANOVA was performed for each type of error. All error types showed a main effect of step (all $F(3,123) > 3.79$, all $ps < 0.03$, all $\eta_p^2 > 0.08$), and linear trend analyses indicated an overall increase in error across steps (all $F(1,41) > 7.70$, all $ps < 0.01$, all $\eta_p^2 > 0.15$). A main effect of search array was found for same episode different step errors, as well as for missed responses (both $F(2,82) > 14.36$; both $ps < 0.001$, both $\eta_p^2 > 0.25$), reflecting higher errors on the first search array of each step. Finally, same episode different step errors showed a significant step × array interaction ($F(6,246) = 5.96$, $p < 0.001$, $\eta_p^2 = 0.13$). A similar ANOVA for reaction time also showed a significant main effect for step ($F(3,123) = 17.21$, $p < 0.001$, $\eta_p^2 = 0.30$), a significant main effect for search array ($F(2,82) = 234.42$, $p < 0.001$, $\eta_p^2 = 0.85$), and a significant step × array interaction ($F(6,246) = 9.83$, $p < 0.001$, $\eta_p^2 = 0.19$).

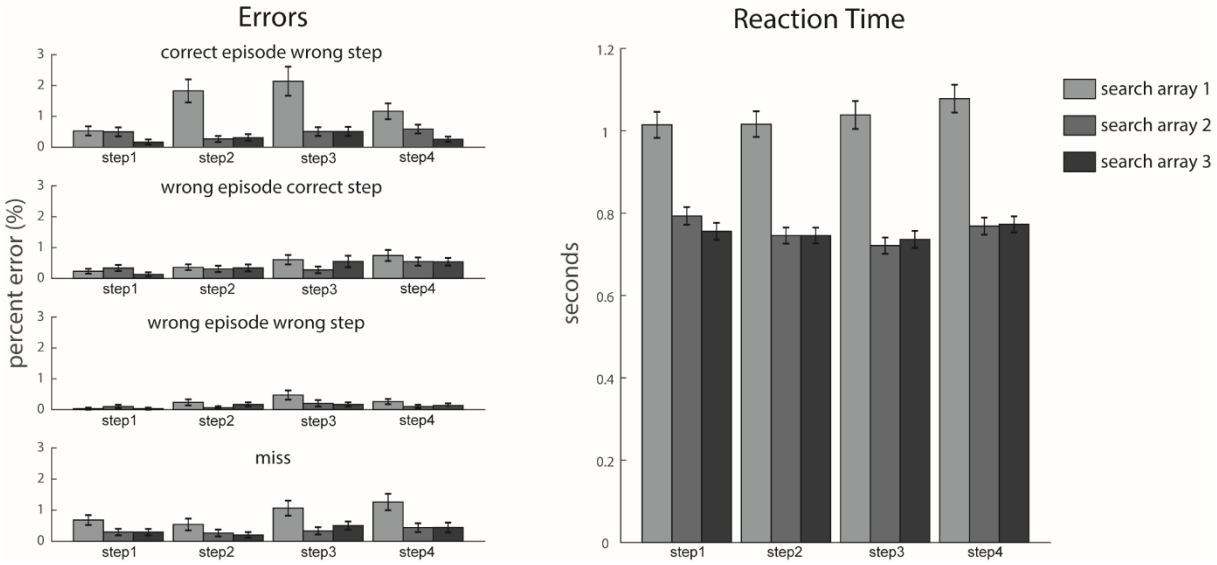


Figure 3.3. Behavioral performance summarized according to four possible error types (choosing an item from same episode different step, different episode same step, different episode different step, and missed response), as well as reaction time for correct trials. For each step, the three bars indicate performance on each of the three successive search arrays. Error bars indicate standard error of the mean.

3.4.2 Univariate results

3.4.2.1 ROI analysis

The FIR model provided estimates of the observed BOLD response timecourse across a pair of task episodes, in successive 1.5 s windows starting from the onset of the first step. In the main analysis, we extracted these FIR responses from *a priori* networks (Figure 3.4) The MD network exhibited positive activity throughout each episode, along with four peaks corresponding to the four steps. These results suggest involvement in setting up and executing individual task steps. Additionally, overall MD activity gradually increased throughout the task episode, suggesting that the MD network is also sensitive to progress through the episode. For DMN regions, in contrast, tonic activation began below baseline but gradually increased through the episode, culminating in a large phasic response at episode completion. For both networks, the signal clearly resets between episodes.

To quantify the phasic and tonic components contributing to the BOLD response at each task step, we performed a complementary GLM analysis with onset and epoch regressors modelling each task step. Four onset regressors were designed to reflect phasic activity at the

onset of each task step. The offset regressor was included to capture the phasic activity at the end of the episode. Finally, epoch regressors were designed to reflect tonic activity throughout each step.

Within the MD network, there were strong onset responses, in line with FIR results. Contrasts with baseline showed that all four step onsets were significantly greater than baseline (all *t*s > 10.91, all *p*s < 0.001, all ds > 1.68) and there was a smaller yet significant offset response (t = 2.48, p = 0.02, d = 0.38). A one-way repeated measures ANOVA showed a significant difference across the four step onsets ($F(3,123) = 5.60$, $p < 0.01$, $\eta_p^2 = 0.12$), with a quadratic ($F(1,41) = 21.61$, $p < 0.001$, $\eta_p^2 = 0.35$) but not linear ($F(1,41) = 0.22$, $p = 0.64$, $\eta_p^2 < 0.01$) trend across steps, reflecting an increasing response across steps 2-4, but a disproportionate response to the onset of the first step, i.e. the onset of the entire episode. Looking at epoch regressors, all four epoch responses were greater than baseline (all *t*s > 3.96, all *ps* < 0.001, all ds > 0.61). ANOVA showed a significant main effect of step ($F3,123) = 7.73$, $p = 0.01$, $\eta_p^2 = 0.16$), as well as a significant linear ($F(1,41) = 9.48$, $p < 0.01$, $\eta_p^2 = 0.19$) and quadratic trend ($F(1,41) = 5.08$, $p = 0.03$, $\eta_p^2 = 0.11$), reflecting an increasing but saturating response.

The DMN network showed a different profile. Only the onset of the first step (*t* = 3.22, *p* < 0.01, d = 0.50) and the offset response at the end of the episode (*t* = 4.38, *p* < 0.001, d = 0.68) were greater than baseline. Step onsets 2-4 were not significantly different from baseline (all |t|s < 2.09, all ps > 0.07, all |d|s < 0.33). ANOVA of the four step onsets showed a significant main effect of step ($F(3,123) = 9.87$, $p < 0.001$, $\eta_p^2 = 0.19$), as well as significant linear ($F(1,41) = 9.70$, $p < 0.01$, $\eta_p^2 = 0.19$) and quadratic ($F(1,41) = 7.16$, $p = 0.01$, $\eta_p^2 = 0.15$) trends, confirming the larger response to the first onset. Among the epoch responses, the first step was significantly lower than baseline (t = -3.21, p = 0.01, d = -0.49; for steps 2-4 all |t|s < 1.60, all ps > 0.23, all |d|s < 0.19). ANOVA showed a significant main effect of step ($F(3,123) = 18.42$, $p < 0.001$, $\eta_p^2 = 0.31$), as well as a significant linear trend ($F(1,41) = 38.89$, $p < 0.001$, $\eta_p^2 = 0.49$), suggesting an increase in activation across steps. As seen in the FIR time-course, this implies a gradual release of tonic deactivation across the duration of the task episode.

To examine whether the profiles of different regions within each network showed unique responses, we performed the same analyses (FIR timecourse modelling and event-based GLM) on individual ROIs (see Appendix B. Figure B.3). Results showed that trends of activation across the four steps for individual ROIs were largely similar to the network in which

they belong, though there were some individual differences between ROIs. The aMFG and AI showed negative epoch responses, in contrast to other MD regions. The PHC showed positive epoched responses, in contrast to other DMN regions.

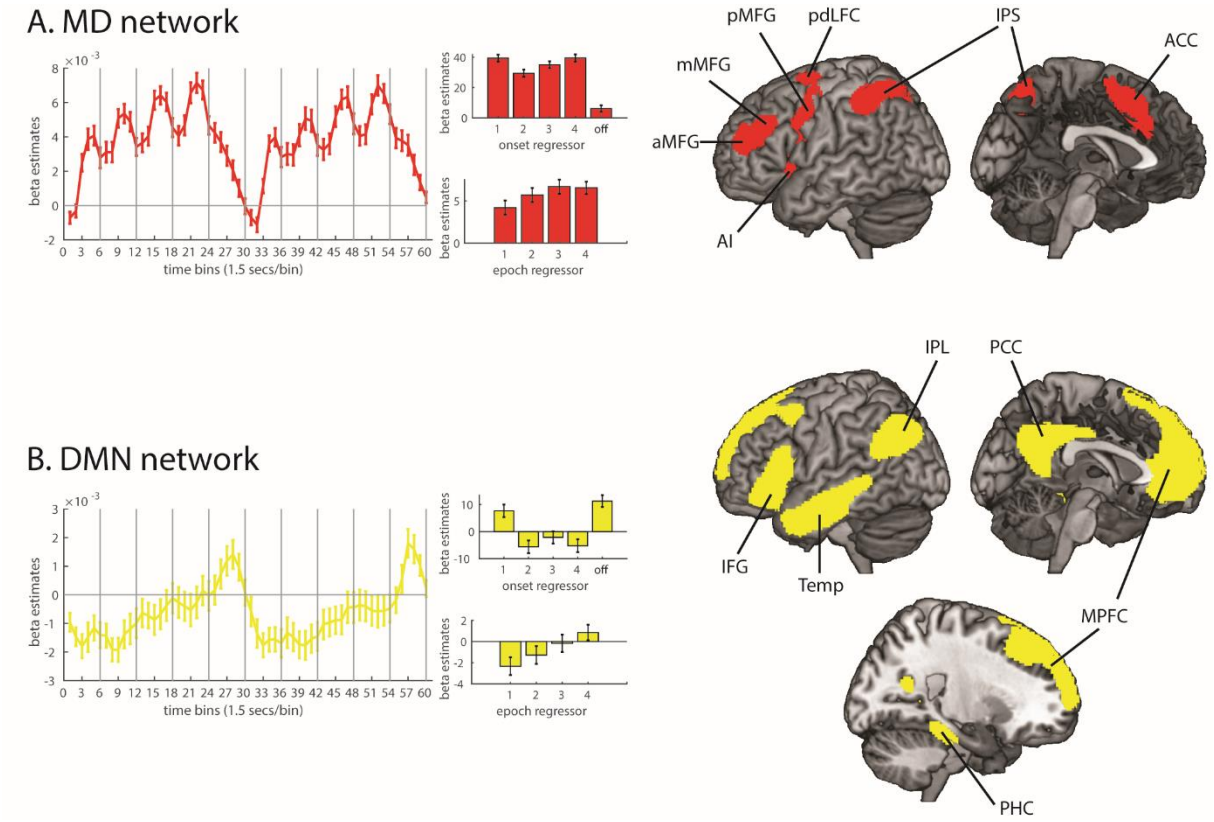


Figure 3.4. Univariate analysis of the MD network and DMN responses across task episodes. The left plot shows results of the FIR analysis, with BOLD response as a function of time as participants progressed across two consecutive episodes. The upper middle plot shows beta estimates associated with the step onsets (bars 1-4), and with the end of the episode (bar 5). The lower middle plot shows beta estimates for epoch regressors for each step. Error bars indicate standard error of the mean. The right panel depicts the location of regions within each network.

3.4.2.2 Whole-brain analysis

Results from the whole-brain analysis, again separating onset and epoch regressors, are shown in Figure 3.5. Panels Ai and Aii show contrasts of average onset and epoch regressors against baseline, with regions in red and blue indicating above and below baseline responses. Panels Bi and Bii show increasing (red) and decreasing (blue) linear trends across task steps.

In comparison to baseline, the mean step onset response (Figure 3.5Ai) was significantly positive in large parts of the MD network, including regions of lateral frontal, insular, dorsomedial frontal, and lateral parietal cortex. A mean step onset response was also seen in visual cortex and subcortical structures including the cerebellum. The mean step onset response was significantly negative in many core DMN regions. Mean epoch responses greater than baseline (Figure 3.5Aii) were more restricted, including parietal and dorsomedial frontal regions overlapping with the MD ROIs, as well as expected regions of visual and motor cortex. Again, we see negative epoch responses in parts of the DMN. We next examined activity changes across steps 1-4. Onset regressors showed a linear increase across successive task steps in a restricted subset of MD regions (Figure 3.5Bi). Linear decreases were extensive, including many parts of the DMN (Figure 3.5Bi; compare Figure 3.4). Epoch regressors showed a linear decrease in visual cortex, but otherwise, an extensive pattern of linear increase across much of the brain (Figure 3.5Bii).

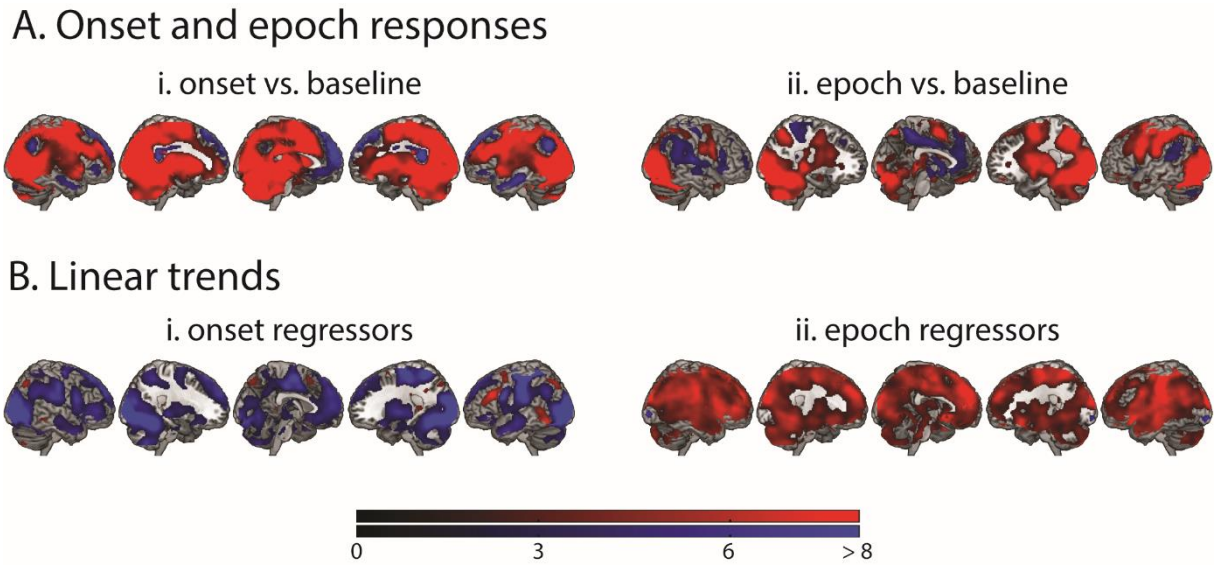


Figure 3.5. Whole brain univariate analysis. (A) shows (i) mean phasic responses to the onset of each step and (ii) mean tonic responses to the duration of each step, versus implicit baseline. Red indicated the positive contrast, while blue indicates the negative contrast. (B) Increasing (red) and decreasing (blue) linear trends of (i) onset and (ii) epoch regressors across step. Colors indicate t-values. All activation maps are thresholded at FDR < 0.05.

We further examined activation in response to transitions from one step to another (Appendix B. Figure B.3). Of particular interest was the onset of the first step (initiation of an

episode) and the offset of the fourth step (completion of an episode). Step 1 onset showed a greater activation than baseline as well as the step 2 onset in many parts of the MD, DMN, and visual cortex. Episode completion responses were also significantly greater than baseline in many brain regions, including parts of both MD and DMN networks; though in MD, the offset response was weaker than the preceding onset response, while in DMN, the offset was significantly higher than the previous onset.

The results may be summarized as follows. Most MD regions showed positive onset and epoch responses to all steps, suggesting direct involvement in setting up and executing task steps, along with visual cortex. Other salient aspects of brain activity concerned the large-scale structure of the episode, including gradually increasing activity as the episode progressed, along with phasic responses at onset and offset of the whole episode. Interestingly, these were apparent in both the DMN and MD network, as well as other brain regions.

3.4.3 RSA results

Results of the RSA analysis are shown in Figure 3.6. Coding of different types of information: room, episode, step, and item are plotted for each network in the left panels. Right panels show whole-brain searchlight results. Analyses of individual ROIs within the two networks are shown in Appendix B. Figure B.4.

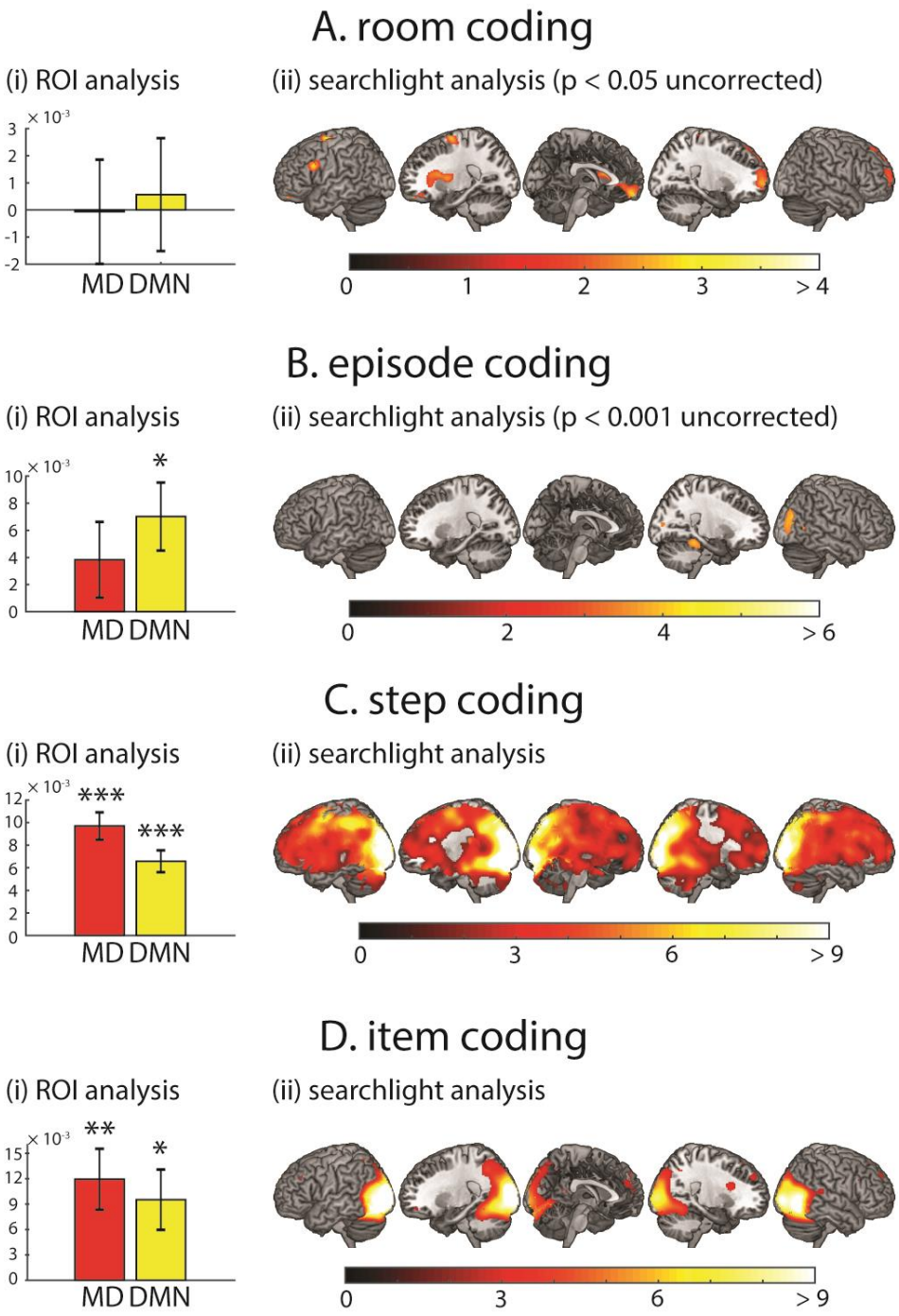


*Figure 3.6. LDC contrasts representing strength of (A) room, (B) episode, (C) step, and (D) item coding in (i) MD and DMN ROIs and (ii) the whole-brain searchlight. Error bars represent standard error. *** indicates $p < 0.001$, ** indicates $p < 0.01$, and * indicates $p < 0.05$ for 1-tailed t-tests against zero. Whole-brain maps for step and item coding are thresholded at FDR < 0.05, whereas room coding is thresholded at $p < 0.05$ (uncorrected), and episode coding is thresholded at $p < 0.001$ (uncorrected).*

Room coding

Neither the MD nor DMN network showed significant room coding (both $|t|$s < 0.04, both ps > 0.78, both $|d|$s < 0.05). Similarly, none of the individual ROIs showed significant room coding (all $|t|$s < 1.86, all ps > 0.79, $|d|$s < 0.29). In the searchlight analysis, we did not find room coding in any region after FDR correction. At the lenient threshold of uncorrected p < 0.05, the most prominent was the MPFC, as well as some regions around the IFG and AI.

Episode coding

The DMN network showed significant coding of episode (t = 2.79, p = 0.02, d = 0.43), while MD did not (t = 1.37, p = 0.18, d = 0.21); however, the difference between networks was not significant (t = 1.55, p = 0.13, d = 0.19). None of the individual ROIs showed significant episode coding after FDR correction for multiple comparisons across ROIs. PCC, PHC, and ACC showed task coding before correction (all ts > 2.05, all ps < 0.05, all ds > 0.31). In the whole-brain searchlight, we did not find any regions that showed episode coding using FDR < 0.05; however, the right PHC and right middle occipital cortex was significant at p < 0.001 uncorrected.

It is possible that the response to regressors modelling adjacent steps could be similar due to imperfect temporal separation of the signal, such that pairs of steps within the same task appear more similar than those from different tasks due to differences in temporal separation in addition to differences in task episode. We examined this possibility by fitting four separate linear regression models using subsets of cells, chosen to differ in separation of zero, one, two, or three steps. That is, we extracted LDC values from cells of the RDM that represented one (step 1 vs. step 2, step 2 vs. step 3, and step 3 vs. step 4), two (steps 1 vs. step 3 and step 2 vs. step 4), or three (step 1 vs. step 4) steps apart, and fitted a model with room, episode, and visual overlap regressors. If temporal leakage were contributing to activity patterns, and hence to apparent episode coding in the DMN, we should expect a stronger effect for steps closer together in time. However, we found no evidence of any difference in episode coding in these four conditions ($F(3,123) = 0.39$, $p = 0.61$, $\eta_p^2 = 0.01$), nor a linear trend as a function of step ($F(1,41) = 0.44$, $p = 0.51$, $\eta_p^2 = 0.01$). Episode coding as a function of step difference is shown in Appendix B. Figure B.5.

Step coding

Step coding was significant in both the MD (t = 7.98, p < 0.001, d = 1.23) and DMN (t = 6.85, p < 0.001, d = 1.06) networks. The MD network showed greater step coding compared

to the DMN ($t = 2.93$, $p < 0.01$, $d = 0.44$). Step coding was also significant in all the individual ROIs (all $ts > 2.49$, all $ps < 0.02$, all $ds > 0.38$). Searchlight analysis showed that step coding was widespread in all regions of the brain. This was not surprising, as in our univariate analysis, we observed significant linear trends across the task episode for much of the brain (see Figure 3.5B; visual cortex showed decreasing activity, while most other regions showed increasing activity).

Item coding

Both MD ($t = 3.31$, $p < 0.01$, $d = 0.51$) and DMN ($t = 2.68$, $p = 0.01$, $d = 0.41$) networks showed significant coding of item. The two networks did not significantly differ in item coding ($t = 1.09$, $p = 0.28$, $d = 0.10$). In the individual ROIs, only IPS showed significant item coding after FDR correction for multiple comparisons across ROIs ($t = 4.02$, $p < 0.01$, $d = 0.62$). In the searchlight analysis, we observed significant item coding predominantly in the visual cortex and IPS.

3.5 Discussion

The present study used fMRI to examine how different cortical networks represent task episodes. Specifically, we focused on the MD and DMN networks. Using FIR analysis to capture the evolution of the BOLD response throughout a multistep episode, we found that MD regions showed positive activity throughout the episode, with separate peaks for successive steps. These results suggest involvement in setting up and executing individual task steps. In contrast, the DMN showed overall deactivation. Along with widespread other regions, both DMN and MD showed sensitivity to the large-scale structure of the episode, with phasic responses to onset and offset, and gradually increasing activity as the episode progressed.

Representational similarity analysis suggested that MD regions showed strong coding of individual items but not the entire task episode, while for DMN, both item and episode coding were significant. Both networks additionally showed strong step coding, which was significantly greater in the MD network than the DMN. The RSA searchlight analysis confirmed differential representation of task features. The content of individual task steps (item coding) was represented in visual cortex and the IPS of the MD network. Task identity was most strongly represented in the PHC and lateral occipital cortex, though visible only at a lenient statistical threshold. A lenient threshold also suggested a hint of room coding in the

MPFC. Step was widely represented across most of the brain, in line with strong changes in univariate activity as the task progressed.

The finding that MD regions are especially sensitive to the identity of a current task step and its specific item content is consistent with prior research. Many previous experiments have shown coding of task-relevant information in MD regions that can rapidly change according to task demands (e.g., Li et al., 2007; Woolgar et al., 2011; Freedman et al., 2001), including radical reorganization between successive task steps (Sigala et al., 2008). fMRI studies show strong MD activity when a subgoal is completed and in transitions from one event to another (Sridharan et al., 2007; Farooqui et al., 2012), with progressively increasing activity as a goal is approached (Farooqui et al., 2012; Desrochers et al., 2018). The pattern of MD activation in our study is consistent with these previous findings. The results suggest that, as a task episode progresses, MD representations are in constant flux, reorganizing to encode the detailed contents of each task step.

Univariate results showed a significant peak of DMN activity at the beginning of each episode, but no significant onset responses to subsequent steps. At the completion of the episode, DMN also exhibited a strong offset response. These findings are consistent with prior reports of transient DMN activation at event boundaries (Ben-Yakov et al., 2013, 2014; Baldassano et al., 2018). Our data show, however, that both onset and offset responses are widespread in the brain. It has been proposed that the mental programs required for carrying out a task are assembled at the beginning of task execution (Schneider and Logan, 2006; Farooqui and Manly, 2018a). It is possible that DMN, along with multiple other brain regions, is involved in long-term memory retrieval for the entire task sequence prior to episode initiation. To some extent, our results match the observation that the DMN has long temporal receptive windows and can code for information accumulated over longer time scales (Hasson et al. 2008; Lerner et al. 2011; Manning et al., 2015).

As a whole, the DMN showed significant coding of whole task episodes, with more similar activity patterns for steps from the same compared to different episodes. Searchlight RSA, using more lenient thresholds, revealed a hint of both episode and room coding in specific DMN sub-regions. Episode coding was found in the right PHC and nearby occipital cortex. The PHC is a key component of the DMN, and has been shown to be involved in representation of "situation models", which are higher-level cognitive representations of relationships between different elements of an episode (Diana et al., 2007; Ranganath, 2010b; Ranganath and Ritchey, 2012; Reagh and Ranganath, 2018). Meanwhile, some hint of room coding was found in the

MPFC. This is consistent with the suggestion that the MPFC is implicated in schema representation, capturing similarities across particular episodes at a higher level (Preston and Eichenbaum, 2013; Ghosh and Gilboa, 2014a; Robin and Moscovitch, 2017).

We also found that the DMN showed significant coding for items. These results show some DMN representation not just for full task episodes, but also for specific contents within the episode. It has been suggested that the hippocampus, a key region in the DMN, is involved in binding items to contextual episodes (O'Reilly and Rudy 2001; Diana et al. 2007; Manning et al., 2015; Hsieh et al., 2014). Although we found both item and episode representations coexisting in the DMN, consistent with a compositional code, this experiment cannot determine whether and where items and episodes might be bound into a conjunctive representation: because items were unique to each task, item-episode conjunctions are indistinguishable from item coding. Disentangling these different forms of co-representation requires designs where the same item appears in different contexts. As well as item-context conjunctions in the hippocampus (Hsieh et al., 2014) such designs have associated various frontal and temporal regions with item-order associations (e.g., Reverberi et al. 2012; Kalm and Norris 2014), and rule-rule compositionality (e.g., Cole et al. 2011).

Both MD and DMN, along with most regions of the brain, tracked progress through the task episode, shown by increasing linear trends in the univariate data and step coding in the RSA analysis. These observations are consistent with previous studies that tracked activity and step representation throughout a task episode in MD (Farooqui et al., 2012; Desrochers et al., 2018) and DMN (Hsieh and Ranganath, 2015) ROIs, but suggest that it might be a much more global property of brain function. While visual cortex showed a decrease in sustained activity over time, which may reflect adaptation to the sensory input (Grill-Spector and Malach, 2001; Grill-Spector et al., 2006), most other cortical regions showed an increase in sustained activity over the episode. As this effect was so widespread, it is difficult to offer a precise interpretation, and different areas may increase for different reasons (Kalm and Norris, 2017). For example, it is possible that increased activations in some regions reflect revision and reconfiguration of control representations that may increase in demand as larger portions of the task are complete (Farooqui et al., 2012; Desrochers et al., 2015, 2016). These activity changes could also reflect gradual assembly of an episode representation (Dumontheil et al., 2011) or accumulation of new information (Hasson et al., 2008; Lerner et al., 2011).

A hierarchical control structure is an organized representation of control elements (Rosenbaum et al., 1983; Schneider and Logan, 2006) with task identity, local entities, and

serial position codes. Our results describe how broad brain networks are involved in the execution of task sequences, with MD and DMN regions exhibiting differential timecourses throughout the episode. The DMN, we suggest, may establish overall cognitive context, representing both individual cognitive operations and their broader context, and perhaps involved in binding them together. These functions may be consistent with representing a "situation model" (Ranganath and Ritchey, 2012). At the same time, the MD system, along with sensory regions, contributes the detailed content of individual cognitive operations. Activity in both networks reflect the broad temporal structure of behavior, with phasic activity at task onset and offset, and gradually increasing activity as the whole sequence of behavior progresses. Acting together, multiple brain regions manage the hierarchical structure of goal-directed behavior.

Chapter 4 The functional convergence and heterogeneity of social, episodic, and self-referential thought in the default mode network

4.1 Abstract

The default mode network (DMN) is engaged in a variety of cognitive settings, including social, semantic, temporal, spatial, and self-related tasks. Andrews-Hanna et al. (2010, 2012) proposed that the DMN consists of three distinct functional-anatomical subsystems – a dorsal medial prefrontal cortex (dMPFC) subsystem that supports social processing and introspection about mental states; a medial temporal lobe (MTL) subsystem that contributes to memory retrieval and construction of mental scenes; and a set of midline core hubs that are involved in processing self-referential information. In male and female participants, we examined activity in the DMN subsystems during six different tasks: (1) theory of mind and (2) moral dilemmas (for social cognition), (3) autobiogrpahical memory and (4) spatial navigation (for memory-based construction/simulation), the (5) self/other adjective judgement (for self-related cognition), and finally, a (6) rest condition compared to a working memory task. At a broad level, we observed similar whole-brain activity maps for the six contrasts, and some response to every contrast in each of the three subsystems. In more detail, both univariate analysis and multivariate activity patterns showed partial functional separation, much of it in close accord with the proposals of separate dMPFC and MTL subsystems, though with less support for common activity across anterior and posterior regions of a midline core. Integrating social, spatial, self-related, and other aspects of a cognitive situation or episode, multiple components of the DMN may work closely together to provide the broad context for current mental activity.

4.2 Introduction

The default mode network (DMN) was originally discovered as a collection of medial prefrontal, lateral temporal, lateral parietal, and posterior medial cortical regions that reliably exhibit enhanced activity during passive rest compared to simple, externally oriented tasks (Shulman et al., 1997; Raichle et al., 2001). Raichle et al. (2001) postulated that the DMN is

involved in cognitive states that are suspended during many attentionally-demanding tasks. A large body of literature has now provided evidence that the DMN supports several aspects of spontaneous and deliberate self-generated thought that transcend the immediate sensory environment (Christoff et al., 2004, 2009; Buckner et al., 2008; Andrews-Hanna, 2012; Andrews-Hanna et al., 2014b). Complementing this strong activity during rest, subsequent work has shown DMN activity across a variety of high-level tasks, including social (Greene and Haidt, 2002; Mars et al., 2012; Molenberghs et al., 2016), semantic (Binder et al., 2009; Humphreys and Lambon Ralph, 2017), episodic (Ranganath and Ritchey, 2012; Rugg and Vilberg, 2013), and self-referential (Kelley et al., 2002) cognition.

One common proposal is that the DMN represents broad features of a cognitive episode, scene or context (Hassabis and Maguire, 2007; Ranganath and Ritchey, 2012; Manning et al., 2014; Baldassano et al., 2017). This episode might be imagined, as in spontaneous mind-wandering or recollection of a previous event, or currently perceived (Ranganath and Ritchey, 2012; Manning et al., 2014; Baldassano et al., 2017). Contextual representations might include spatial, social, temporal, self-related and other features, with reduced processing of these features during focused attention on the details of a simple task, but enhancement during spontaneous, self-generated cognition at rest.

A core question is the degree of heterogeneity across DMN regions. Early reviews (Buckner and Carroll, 2007; Buckner et al., 2008), meta-analyses (Spreng et al., 2009; Andrews-Hanna et al., 2014b), as well as experimental data (Spreng and Grady, 2010; Axelrod et al., 2017) suggested that spatial, social, memory and imagination tasks produce substantially overlapping DMN activity. More recently, consistent with the multiple features of a cognitive context, some studies suggest that the DMN exhibits heterogeneous functional components (Andrews-Hanna et al., 2010b, 2014a; Andrews-Hanna, 2012). In an important synthsis, Andrew-Hanna et al. (2010b) partitioned the DMN into three subsystems. A dorsal medial prefrontal cortex (dMPFC) subsystem, composed of the dorsal medial prefrontal cortex (dMPFC), the temporoparietal junction (TJP), the lateral temporal cortex (LTC), and the temporal pole (TempP), is involved in "introspection about mental states", including theory of mind, moral decision making, social reasoning, story comprehension, and conceptual processing. A medial temporal lobe (MTL) subsystem, consisting of the ventromedial prefrontal cortex (vMPFC), the posterior inferior parietal lobe (pIPL), the retrosplenial cortex (RSC), the parahippocampal cortex (PHC), and the hippocampal formation (HF+). subserves "memory-based construction/simulation", including autobiographical memory, episodic future

thinking, contextual retrieval, imagery, and navigation. These two are proposed to converge on a midline core, consisting of the anterior prefrontal cortex (aMPFC) and the posterior cingulate cortex (PCC). The core subserves valuation of "personally significant information", as well as linking social and mnemonic processes shared with the dMPFC and MTL subsystems.

The current study further investigates separation and integration across the DMN. To this end we examined patterns of univariate and multivoxel activity across six tasks, aiming to separate social cognition, memory-based construction/simulation, self-related cognition and rest. Across this combination of tasks and analysis methods, we found a degree of functional separation between DMN regions, largely consistent with the Andrews-Hanna (2010b) dMPFC and MTL subsystems, though less so with their concept of the midline core. To a degree, however, we also found overlapping activity across the whole DMN, with each task producing some activation in each subsystem. While subsystems of the DMN system appear somewhat specialized, our data also suggest collaboration in assembling the multiple components of a cognitive situation or context.

4.3 Methods

4.3.1 Participants

27 participants (13 male, 14 female; ages 20-39, mean = 24.8, SD = 4.3) were included in the experiment at the MRC Cognition and Brain Sciences Unit. An additional participant was excluded due to excessive head motion (> 5 mm). All participants were fluent English speakers, neurologically healthy, right-handed, with normal or corrected-to-normal vision. Participants were also required to be familiar with navigating in Cambridge city centre. Procedures were carried out in accordance with ethical approval obtained from the Cambridge Psychology Research Ethics Committee, and participants provided written, informed consent before the start of the experiment.

4.3.2 Stimuli and task procedures

This study consisted of six tasks that were previously found to engage the DMN. These tasks were: a theory of mind task, a moral dilemmas task, an autobiographical memory task, a spatial imagery task, a self/other adjective judgement task, and a comparison of rest with working memory (Figure 4.1). For the first five tasks, each run contained two conditions (one condition that has been associated with DMN activity and a matched control condition), along

with periods of fixation between trials or blocks. Conditions were presented in randomized order, with the restriction of a maximum of two consecutive trials or blocks of the same condition. For the working memory task, each run contained alternating periods of working memory and periods of fixation. In all runs, participants were instructed to relax and clear their minds of any thought during fixation periods, and fixation periods were jittered and sampled from a random uniform distribution (see details below for each task). Before entering the scanner, participants practiced a shortened version of each task (containing 1~2 trials or blocks of each condition). Participants were also asked to practice writing down digital numbers until they were able to write all of them in the correct format, and to clarify that they were familiar with all 20 landmark locations used in the spatial imagery task. Inside the scanner, there were two scanning runs for each task. Run order was randomized with the constraint that repeats of the same task were between four and seven runs apart. Before the start of each run, participants were played audio-recorded task instructions to remind them of what to do during that run. Each run lasted approximately 5~7 minutes.

All tasks were coded and presented using the Psychophysics Toolbox (Brainard, 1997) in Matlab 2014a (The MathWorks, Inc.). Stimuli were projected on a 1920 × 1080 screen inside the scanner, and participants indicated their responses using a button box, with one finger from each hand in tasks that had two-choice decisions (all tasks except autobiographical memory).

According to Andrews-Hanna (2012), the chosen tasks would be hypothesized to differently engage the dMPFC and MTL subsystems, with all tasks engaging the core hubs. The theory of mind and moral dilemmas tasks were chosen as tasks requiring “introspection about mental states” and were hypothesized to recruit the dMPFC subsystem. The autobiographical memory and spatial imagery tasks were chosen as tasks that required “memory-based construction/simulation” and were hypothesized to recruit the MTL subsystem. The self/other judgement task was chosen as a task that involved “personally significant information”, and was hypothesized to recruit predominantly the core hubs. Finally, the working-memory task was chosen to examine the activity of the DMN during passive rest compared to an external task.

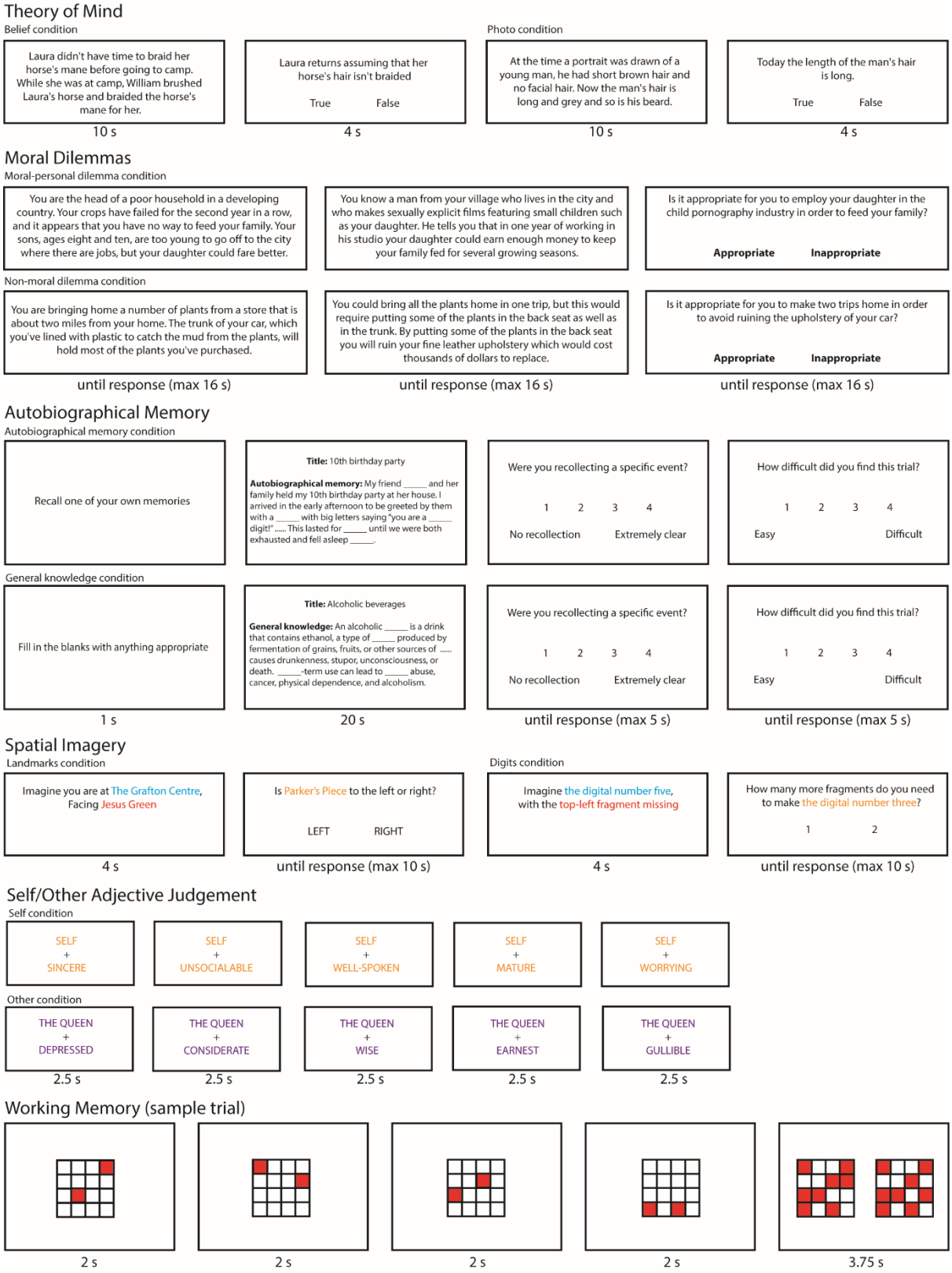


Figure 4.1. Example trial/block from each of the six tasks: theory of mind, moral dilemmas, autobiographical memory, spatial imagery, self/other adjective judgement, and working memory. All stimuli were shown on a 1920 × 1080 screen (stimulus size and width/height ratio

has been adjusted for this figure for illustration purposes). The examples in the autobiographical memory task were shortened to fit in the figure.

4.3.2.1 Theory of Mind task

The theory of mind task was adapted from Dodell-Feder et al. (2010). On each trial, participants were presented with a short story to read for 10 seconds. Afterwards participants were given a statement related to the story and were asked to judge whether it was 'true' or 'false' by pressing a button (left or right). Some trials involved making judgements about other people's beliefs, while others involved making non-belief judgements. Each question stayed on the screen up to ten seconds, or until the participant made a button press. This was then followed by a 10~24 second fixation period before the next trial began. Each run consisted of five trials of each condition (belief and non-belief).

4.3.2.2 Moral Dilemmas task

The moral dilemmas task was adapted from Greene et al. (2001). On every trial, participants were presented with a hypothetical situation that posed a dilemma, which could either be a moral-personal dilemma or a non-moral dilemma. Each dilemma was presented as text through a series of three displays, with the first two describing a scenario and the third posing a question about the appropriateness of an action one might perform in such a situation. The maximum time one display could be on screen was 16 seconds, but when participants finished reading the text, they were allowed to press any button to move on to the next display. On the third display, participants made the appropriateness judgement by pressing a button (left or right). They were told that there was no correct answer for many of the questions, and were asked to consider each situation carefully and provide their best answer. A 6~8 second fixation cross was presented in between each trial. Each run consisted of 5 trials of each condition (moral-personal dilemma (MPD) or a non-moral dilemma (NMD)).

4.3.2.3 Autobiographical Memory task

Prior to the experiment, participants were asked to provide 12 written personal memories, each with a title that provided a general description of its contents. Participants were given specific instructions to provide clear memories, where they were able to remember the people, objects, and location details featured in the corresponding memory. Each memory was required to be between 100~150 words long. All events were required to be temporally and contextually specific, occurring over minutes or hours, but not more than one day. The

memories were then edited by the experimenter such that 13~17 critical words were removed and replaced with a blank underscore line. Occasionally, if the memories participants sent were too long, they were shortened; or if the memories were too short or contained too few details, a new sentence with a prompt was added (e.g., “I was wearing a _____”, “It was around _____ o’clock”, “I felt very _____”).

During the task, on a given trial, participants were given a 100~150-word long text to read with 13~17 critical words missing, and were asked to fill in the blanks in their mind. Half of the trials used text adapted from the participants’ autobiographical memories; the other half contained text related to general knowledge (either procedural tasks, such as “how to make chocolate chip cookies”, or knowledge about a common topic, such as “alcoholic beverages”). Before the onset of the text display, a 1 second cue was presented to indicate the upcoming condition. For autobiographical memory trials, participants were told to try to “really get into the memory” while filling in the blanks. They were asked to try to imagine themselves reliving that experience. In the general knowledge condition, participants were asked to fill in the blanks with anything appropriate, and to try to “think carefully for good answers”. All trials were terminated after 20 seconds. However, participants were told that there was no need to rush to try to finish all the blanks, and it was more important to be accurate than fast. This was designed to encourage participants to be engaged as much as possible throughout the 20 seconds. After the 20 seconds were over, participants were given two rating questions ('Were you recollecting a specific event?' and 'How difficult did you find this trial?'). They were given five seconds to provide each rating on a scale of 1 to 4. Since it involved four buttons, participants gave responses with the four fingers of their right hand. This was then followed by an 8~12 second fixation period between trials. There were five trials of each condition (autobiographical memory and general knowledge) in each run.

4.3.2.4 Spatial Imagery task

In the spatial imagery task, there were two types of mental imagery conditions, each presented in blocks of trials. One type of block involved judging relative locations of landmarks in Cambridge (this task was adapted from Vass & Epstein, 2017). On each trial, there was first a four second instruction to imagine standing at the landmark indicated in the first line (e.g., Botanic Garden) while facing the landmark indicated in the second line (e.g., King’s College). Afterwards, participants were shown a second screen with a new landmark location (e.g., Parker’s Piece), and were asked to indicate whether it would be on their left or right (in this example the correct answer would be right). The question stayed on the screen for up to 10

seconds, or until participants made a button press. The other type of block involved judging how many fragments were needed to complete a target digital number. At the beginning of each trial, a four second instruction was given to imagine a digital number indicated in the first line (e.g., three) with either an additional fragment or a fragment missing indicated in the second line (e.g., top-right fragment missing). Afterwards, participants were shown a new screen indicating a new target digit (e.g., five), and were asked how many more fragments would need to be added to their original mental image to complete the target (in this example the correct answer would be 1). Participants had up to 10 seconds to answer 1 or 2 (left and right buttons). The two conditions (landmarks and digits) were presented in blocks of four trials, with a 6~16 second fixation period in between each block. There were four blocks of each condition per run.

4.3.2.5 Self/Other Adjective Judgement task

The self/other judgement task was adapted from Kelley et al. (2002). A total of 160 adjectives were selected from a pool of normalized personality trait adjectives (Anderson, 1968). Half of the words were positive traits and half were negative. On each trial, participants were asked to make a yes/no judgement via button press to indicate whether an adjective shown on the bottom of the screen described the person indicated on the top of the screen (self or the Queen). Each trial was presented for a fixed period of two seconds followed by a 0.5 second fixation. The task was grouped into blocks according to "self" and "the Queen", with each block consisting of five trials. There were eight blocks of each condition per run. A 6~16 second fixation period separated each block.

4.3.2.6 Working Memory task

The working memory task was adapted from Fedorenko et al. (2013). On each trial, participants were presented with four consecutive displays. Each display was a 4 × 4 grid, with two of the cells colored red and the remaining white. Each display was presented for two seconds. Afterwards, participants were presented two choice displays, on the left and right of the screen, one of which had eight red cells in locations corresponding to those from the previous four displays, while the other was similar but with one cell misplaced. Participants were given 3.75 seconds to indicate the correct display by pressing left or right. This was followed by a 0.25 second feedback on the accuracy of their choice. There was a 12~16 second fixation period between trials. Each run consisted of 16 trials.

4.3.3 fMRI data acquisition and preprocessing

Scanning took place in a 3T Siemens Prisma scanner with a 32-channel head coil. Functional images were acquired using a standard gradient-echo echo-planar imaging (EPI) pulse sequence (TR = 2000 ms, TE = 30 ms, flip angle = 78°, 64 × 64 matrices, slice thickness = 3 mm, 25% slice gap, voxel size 3 mm × 3 mm × 3 mm, 32 axial slices covering the entire brain). The first five volumes served as dummy scans and were discarded to avoid T1 equilibrium effects. Field maps were collected at the end of the experiment (TR = 400 ms, TE = 5.19 ms / 7.65 ms, flip angle = 60°, 64 × 64 matrices, slice thickness = 3 mm, 25% gap, resolution 3 mm isotropic, 32 axial slices). High-resolution anatomical T1-weighted images were acquired for each participant using a 3D MPRAGE sequence (192 axial slices, TR = 2250 ms, TI = 900 ms, TE = 2.99 ms, flip angle = 9°, field of view = 256 mm × 240 mm × 160 mm, matrix dimensions = 256 × 240 × 160, 1 mm isotropic resolution).

The data were preprocessed and analyzed using automatic analysis (aa) pipelines and modules (Cusack et al. 2014), which called relevant functions from Statistical Parametric Mapping software (SPM 12, http://www.fil.ion.ucl.ac.uk/spm) implemented in Matlab (The MathWorks, Inc., Natick, MA, USA). EPI images were realigned to correct for head motion using rigid-body transformation, unwarped based on the field maps to correct for voxel displacement due to magnetic-field inhomogeneity, and slice time corrected. The T1 image was coregistered to the mean EPI, and then coregistered and normalized to the MNI template. The normalization parameters of the T1 image were applied to all functional volumes. Spatial smoothing of 10 mm FWHM was applied for whole-brain univariate second-level analysis, but no smoothing was applied for ROI-based analyses or multi-voxel pattern analysis.

A general linear model (GLM) was estimated per participant and per voxel for each of the six tasks. A high-pass filter with 1/128Hz cutoff was applied to both the data and the model. For the first five tasks, regressors were created for each condition, with fixation periods serving as implicit baseline. In the working memory task, one regressor was created for the fixation periods to model passive fixation as the contrast against active task as implicit baseline. Error trials (only applicable for the theory of mind and spatial imagery tasks) and no-response trials were modelled using a separate regressor. All regressors were created by convolving the interval between stimulus onset and response (or display offset when no responses were required) with the canonical hemodynamic response function. Run means and movement parameters were included as covariates of no interest. The resulting beta-estimates were used

to construct contrasts between the two conditions of each task, or for working memory, the contrast of rest against task as implicit baseline.

4.3.4 Whole-brain univariate analysis

The between-condition contrasts that were used to examine DMN activity were: (1) belief > non-belief in the theory of mind task; (2) moral-personal dilemmas > non-moral dilemmas in the moral dilemmas task; (3) autobiographical memory > general knowledge in the autobiographical memory task; (4) landmarks > digits in the spatial imagery task; (5) self > other in the self/other adjective judgement task; and (6) rest > working memory.

A second level whole-brain analysis (one-sample t-test across subjects) was conducted on each of the six within-subject contrasts above, to obtain group activation maps for each contrast separately. Activation maps were thresholded at $p < 0.05$, controlling the false discovery rate (FDR; Benjamini and Yekutieli, 2001). A whole-brain analysis was conducted to examine individual participant activations for each of the 6 contrasts. For each voxel, we computed the number of participants with significant activation, applying FDR correction across all voxels of all participants (Heller et al., 2007). This resulted in a whole-brain map showing the number of participants with significant activation within each voxel. Based on the 6 random-effects analyses above, a similar map was constructed to show the number of significant task contrasts at each voxel (Heller et al., 2007). MRIcroN (Rorden et al., 2007) was used for visualization of whole-brain maps.

4.3.5 Regions of interest and ROI analysis

A DMN mask was constructed using the 17 network parcellation from Yeo et al. (2011), concatenating networks 10, 15, 16, and 17. Networks 15, 16, and 17 largely corresponded to the three DMN subnetworks described in Andrews-Hanna (2012), which are the MTL subsystem, the dMPFC subsystem, and the core hubs. Network 10 was described in Yeo et al. (2011) as the orbital frontal-temporopolar network, which consists of temporopolar and orbital frontal regions. This network was added to the three DMN networks from Yeo et al. (2011) to include the vMPFC region described by Andrews-Hanna (2011). To create a single symmetrical volume, ROI masks (1 for voxels within the region; 0 outside) from the left and right hemispheres were combined using a logical OR operation, then projected back to both hemispheres. The combined network was then slightly smoothed (4 mm FWHM), and voxels with values > 0.5 after smoothing were retained. Finally, the combined network was parcellated into 20 smaller subregions by assigning each voxel to its closest DMN coordinate described by

Andrews-Hanna et al. (2010b). The coordinates are listed in Table 1. In cases where non-contiguous volumes were assigned to the same region, any volumes of < 45 voxels were discarded, and the remaining volume with center of mass closest to the Andrews-Hanna coordinate was chosen. The resulting ROIs are shown in Figure 4.2.

Table 4.1. Peak DMN coordinates described in Andrews-Hanna et al. (2010b). Coordinates are based on the Montreal Neurological Institute coordinate system.

Region	Abbreviation	x	y	z
dMPFC subsystem				
Dorsal medial prefrontal cortex	dMPFC	0	52	26
Temporal parietal junction	TPJ	-/+54	-54	28
Lateral temporal cortex	LTC	-/+60	-24	-18
Temporal pole	TempP	-50/+50	14	-40
MTL subsystem				
Ventral medial prefrontal cortex	vMPFC	0	26	-18
Posterior inferior parietal lobule	pIPL	-/+44	-74	32
Retrosplenial cortex	Rsp	-/+14	-52	8
Parahippocampal cortex	PHC	-/+28	-40	-12
Hippocampal formation	HF+	-/+22	-20	-26
Core hubs				
Anterior medial prefrontal cortex	aMPFC	-/+6	52	-2
Posterior cingulate cortex	PCC	-/+8	-56	26

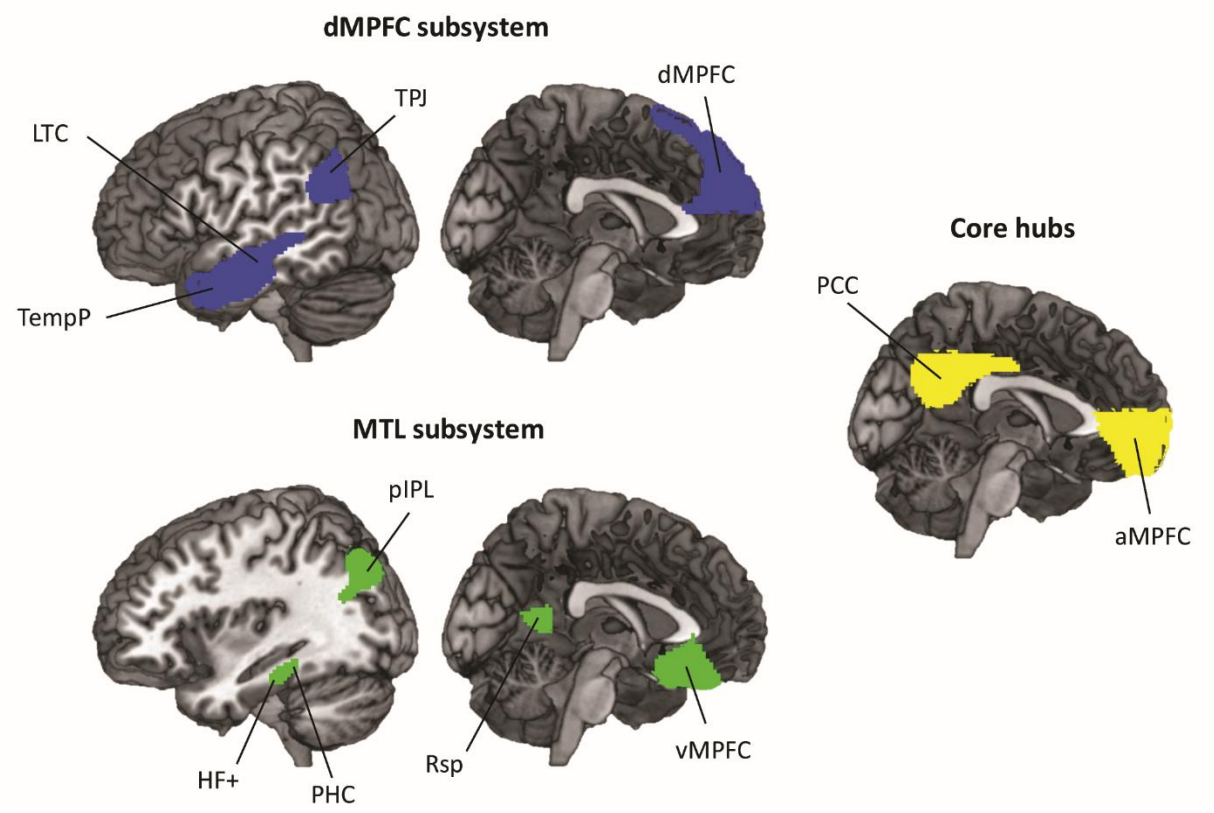


Figure 4.2. DMN ROIs used the in the current experiment. The ROIs are derived from networks 10, 15, 16, 17 described in the 17 network parcellation in Yeo et al. (2011) and devided according to coordinates described in Andrews-Hanna et al. (2010). Regions in blue are part of the dMPFC subsystem, and include the midline dMPFC and bilateral TPJ, LTC, and TempP. Regions in green are part of the MTL subsystem, and include the midline vMPFC and bilateral pIPL, Rsp, PHC, and HF+. The core hubs are represented in yellow, and include the bilateral aMPFC and PCC. For abbreviations see Table 4.1.

For each task, the contrast between the two conditions was averaged within each ROI using the MarsBAR toolbox (Brett et al., 2002). For working memory, the relevant contrast was simply rest against implicit baseline (active task). Contrasts were tested against zero using two-tailed t-tests across subjects, corrected using FDR < 0.05 for multiple comparisons across ROIs. ROI x task ANOVAs were used to examine differences in ROI activity across different contrasts. Finally, the vector of contrast values from all tasks (six in total) was compared across ROIs. Distances between activation profiles for each pair of ROIs were calculated using 1 - Pearson's r, and classical multidimensional scaling (MDS) was used to visualize the differences in activation pattern between ROIs as 2-dimensional distances.

4.3.6 Task-wise multi-voxel pattern similarity

For each ROI, we wished to examine similarity of voxelwise activity patterns across the six tasks. For each participant, we extracted the beta-values for each contrast for each task, and compared the multivoxel patterns of these values between tasks. The similarity between each pair of tasks was measured by Pearson's r, producing a symmetrical 6 x 6 matrix of similarities for each ROI. For each ROI, we quantified which regions showed (1) greater pattern similarity between the two tasks that required "introspection about mental states" (theory of mind and moral dilemmas), compared to similarity of these tasks to others, (2) greater pattern similarity between the two tasks that required "memory-based construction/simulation" (autobiographical memory and spatial imagery), compared to similarity of these tasks to others, (3) a relatively unique pattern for the self/other judgement task (greater similarity for task pairs not including self/other), and (4) a relatively unique pattern for rest (greater similarity for task pairs not including rest). To do this, we created four model similarity matrices based on these *a priori* groupings and evaluated fits to each ROI's task similarity matrix using Kendall's tau-a for each subject, as recommended when the model similarity matrix has ties (Nili et al., 2014). Correlations were tested against zero using 2-tailed t-tests across subjects, and all tests were corrected for multiple comparisons (FDR < 0.05) across the number of ROIs.

To compare patterns of task similarities between ROIs, we used vectors of between-task correlation from the above analysis (15 between-task correlations for each ROI). Similarly to the univariate analysis, distances between each pair of ROIs were calculated using 1 minus the correlation (Pearson's r) between these vectors. Again, multidimensional scaling (MDS) was used for visualization.

4.4 Results

4.4.1 Behavioral results

Mean reaction times (RT) for all responses are summarized in Table 4.2. The first three subjects' RTs for the working memory task were not recorded due to technical error and were excluded in the analysis. Mean accuracies for the theory of mind, mental imagery, and working memory tasks are also summarized in Table 4.2, along with mean ratings of recollection and difficulty for the autobiographical memory task.

Paired t-tests were conducted between the two conditions of the first five tasks, with no correction for multiple comparisons, to examine how well-matched each of the two conditions were within a task. There were no differences in reaction time between the pairs of conditions in the theory of mind, moral dilemmas, autobiographical memory, and self/other adjective judgement task (all |t|s < 1.45, all ps >= 0.16). In the spatial imagery task, RTs were shorter for the landmarks condition than for the digits condition (t = -2.74, p = 0.01). There were no differences in accuracy between the pairs of conditions in the theory of mind and spatial imagery task (both |t|s < 1.62, both ps >= 0.12). As expected, ratings of recollection were significantly greater in the autobiographical memory condition than in the general knowledge condition (t = 21.01, p < 0.001); autobiographical memory was also rated less difficult than general knowledge (t = -4.47, p = 0.001).

Table 4.2. Reaction times (RT), accuracies, and ratings of each condition (mean ± standard error).

	Theory of Mind		Moral Dilemmas		Autobiographical memory		Spatial imagery		Self/Other Adjective Judgement		Working memory
	Belief	Non-belief	MPD	NMD	Memory	Knowledge	Landmarks	Digits	Self	Other	Working memory
RT (s)	3.24 ± 0.03	3.10 ± 0.02	3.50 ± 0.03	3.26 ± 0.04	1.84 ± 0.02	1.90 ± 0.02	2.30 ± 0.02	3.02 ± 0.05	1.44 ± 0.01	1.47 ± 0.01	1.84 ± 0.02
Accuracy (% correct)	87.4 ± 6.2	91.1 ± 4.7	N/A	N/A	N/A	N/A	91.7 ± 5.7	84.3 ± 8.6	N/A	N/A	76.4 ± 5.0
Rating	N/A	N/A	N/A	N/A	Recollection 3.65 ± 0.02 Difficulty 1.41 ± 0.02	Recollection 1.36 ± 0.01 Difficulty 1.91 ± 0.02	N/A	N/A	N/A	N/A	N/A

4.4.2 Whole-brain univariate analysis

A whole-brain random effects analysis was conducted separately for each of the six contrasts of interest (Figure 4.3A; belief > non-belief; moral-personal dilemmas > non-moral dilemmas; autobiographical memory > general knowledge; landmarks > digits; self > other; and rest > task). Consistent with previous findings, the group analysis revealed many regions

that are commonly associated with the DMN. In most tasks, we see activation in the medial prefrontal cortex (MPFC) and posterior medial cortex including PCC, precuneus, and Rsp, as well as temporal and parietal regions on the lateral surface, including pIPL, TPJ, and LTC. Activity for the self/other adjective judgement task was less typical of the DMN, though strongly activated a large portion of the MPFC.

To further quantify consistency across subjects, we computed a whole-brain overlay map for each task, where warmer colors indicate greater number of participants with significant activations (Figure 4.3B). The subject overlay map is largely consistent with the random effects results, as expected, but also indicates variability across participants.

Next, we identified regions that were consistently significantly activated across multiple contrasts (Figure 4.3C). No region was found to be active in all six contrasts after correcting for multiple comparisons (FDR < 0.05). However, several regions showed significant involvement in at least five contrasts. These include the MPFC (including dMPFC, aMPFC, and vMPFC), PCC, pIPL, TPJ, and parts of the LTC.

The results show that all six manipulations activated much of the DMN, and in particular, voxels within the MPFC, PCC, pIPL, TPJ, and LTC were significantly active for at least five manipulations. The theory of mind and moral dilemmas tasks showed strong activation of dMPFC, while the autobiographical memory and spatial imagery tasks showed peaks in vMPFC. These differences correspond to Andrews-Hanna's (2012) observation of the dMPFC being involved in "introspection about mental states" and the vMPFC being involved in "memory-based construction/simulation". Furthermore, the theory of mind and moral dilemmas tasks activated more anterior portions of the IPL than the autobiographical memory and spatial imagery tasks. This again corresponds to the separation of the TPJ (more anterior) and pIPL (more posterior) regions of the IPL, and matches their assignment to the dMPFC and MTL subsystems. The self > other contrast most consistently activated the MPFC across subjects, one of the core hubs identified by Andrews-Hanna (2012) to be responsive to "personally significant information". However, the other hub region, the PCC, was only weakly activated. Our results show activity across much of the DMN for multiple contrasts, along with a degree of differentiation between dMPFC and MTL subsystems.

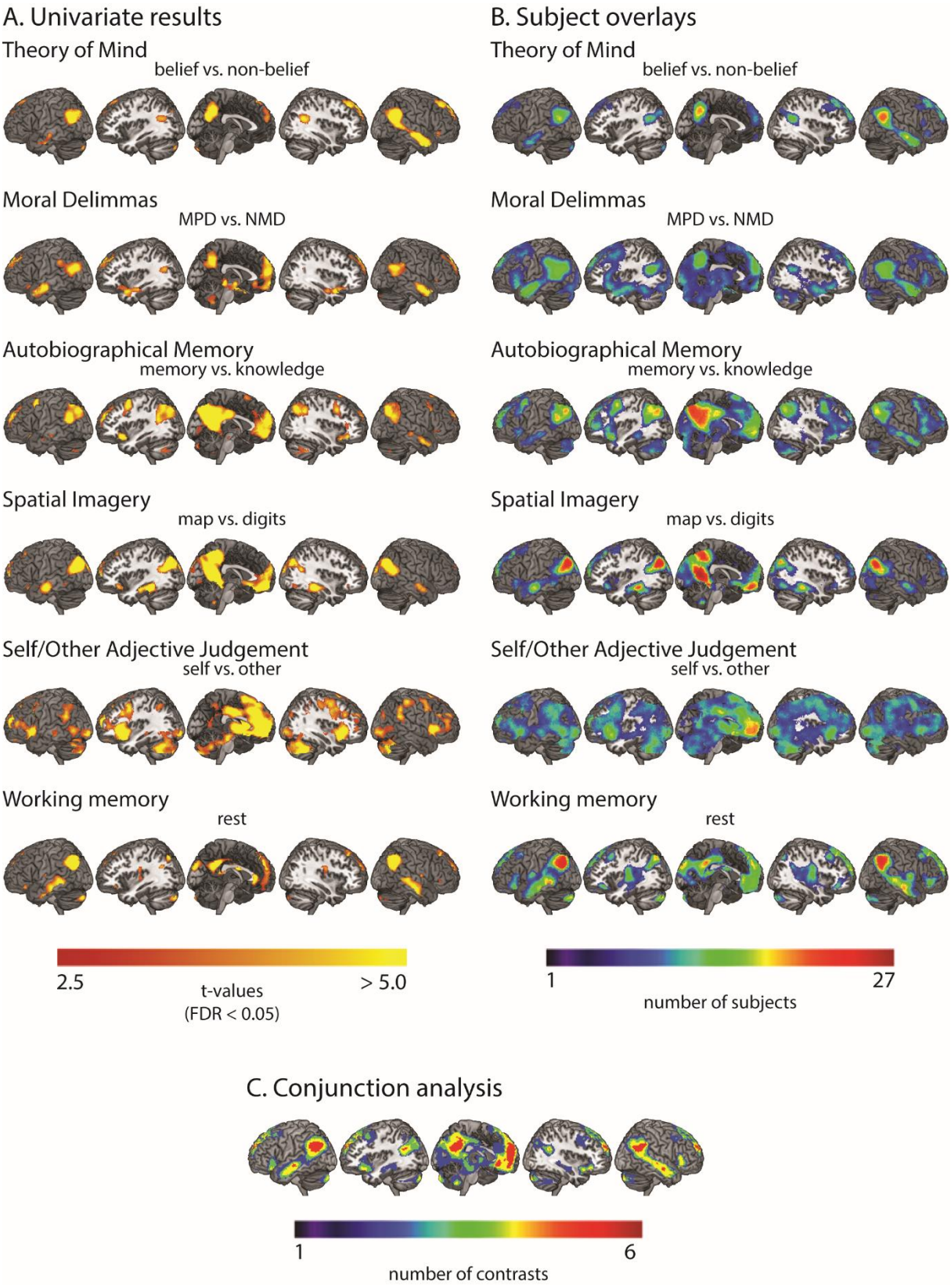


Figure 4.3. Univariate activity showing recruitment of the DMN network by all six tasks. (A) Whole-brain t-maps of the contrasts of interest in the six tasks. This includes belief > non-belief in the theory of mind task; moral-personal dilemmas > non-moral dilemmas in the moral dilemmas task; autobiographical memory > general knowledge in the autobiographical memory task; landmarks >

digits in the spatial imagery task; self > other in the self/other adjective judgement task; and rest > task in the working memory task (working memory as implicit baseline). t-maps were thresholded at $p < 0.05$ (FDR corrected). (B) Overlay map of significant activations found in single subjects in the contrasts of interest. The color of each voxel represents the number of subjects that had significant activation in that voxel for a particular contrast, thresholded at 1 subject. (C) Overlay map of the number of significant contrasts from the six second-level analyses. The color of each voxel represents the number of contrasts that had significant activation in that voxel, thresholded at 2 contrasts.

4.4.3 ROI analysis of univariate activation level

For each of our six contrasts, profiles of activity across DMN ROIs are shown in Figure 4.4A(1). All contrasts were compared against zero using t-tests and were corrected for multiple comparisons with FDR < 0.05.

Examined in detail, profiles suggest some of the anticipated differences between DMN regions, but also some surprises. As expected, theory of mind and moral dilemmas showed significant activation in most regions of the dMPFC and core networks. Activations were also seen in some regions of the MTL subsystem, however, including vMPFC, pIPL and PHC. Averaged contrasts within each network (Figure 4.4A(2)) showed significant activation just for the dMPFC subsystem and core. As anticipated, autobiographical memory and spatial imagery showed strong activations in the MTL subsystem, especially Rsp, and again in the core hubs, but significant activations were also seen in most dMPFC regions. Averaged within subsystems, the response of dMPFC was significantly lower than the other subsystems, but significantly greater than zero. For self-other, activations were more restricted, but included all three regions of the MPFC. Averaged within networks, this contrast was significant in the core and dMPFC subsystem, and, again as anticipated, strongest in the core subsystem. Unlike the previous four contrasts, core activation for self/other was stronger in aMPFC than in PCC.

Perhaps surprisingly, the contrast of rest with working memory showed rather weak activations, significant only in the core and dMPFC subsystem, and significantly negative for some regions of the MTL network. Overall, these results provide broad support for the division into three subsystems, with the dMPFC subsystem especially involved in "introspection about mental states", the MTL subsystem especially involved in "memory-based construction/simulation", and the core hubs involved in all tasks but with particular sensitivity to "personally significant information". At the same time, the results show that separation of networks is far from complete, with at least part of each network activated by every contrast.

Within each network, there are also some clear variations in response. Notably, although the MTL subsystem as a whole was only active in the autobiographical memory and spatial imagery tasks, the pIPL and vMPFC were active for five of the six tasks, similar to the core hubs and dMPFC subsystem.

To compare profiles statistically, the data were entered into a repeated measures ROI (20) × task (6) ANOVA. Consistent with the different profiles suggested by Figure 4.4A(1), there was a strong interaction between ROI and task ($F(95,2470) = 55.57$, $p < 0.001$). There were also significant main effects for task ($F(5,130) = 46.66$, $p < 0.001$) and ROI ($F(19,494) = 25.61$, $p < 0.001$). The interaction in part reflects differences between the three subsystems, so we next repeated the ANOVA using the subsystem average profiles shown in Figure 4.4A(2). The significant interaction ($F(10,260)=100.05$, $p < 0.001$) confirms that this subnetwork grouping captures different functional profiles across the tasks. There were also main effects for networks ($F(2,52) = 15.09$, $p < 0.001$) and task ($F(5,130) = 35.01$, $p < 0.001$). We also wished to test for possible heterogeneity within each subsystem. To this end, ROI × task ANOVAs were repeated for each network separately. For the dMPFC subsystem, there was a significant interaction between ROI and task ($F(30,780) = 9.21$, $p < 0.001$), as well as main effects for ROI ($F(6,156) = 27.54$, $p < 0.001$) and task ($F(5,130) = 12.50$, $p < 0.001$). For the MTL subsystem, we also observed a significant interaction between ROI and task ($F(40,1000) = 34.15$, $p < 0.001$), as well as main effects for ROI ($F(8,208) = 29.78$, $p < 0.001$) and task ($F(5,130) = 110.86$, $p < 0.001$). Finally, there was also a significant interaction ($F(15,390) = 22.92$, $p < 0.001$) as well as main effects of ROI ($F(3,78) = 25.42$, $p < 0.001$) and task ($F(5,130) = 12.87$, $p < 0.001$) in the core hubs.

The distance matrix (Figure 4.4B), based on the dissimilarity of activation profiles for the 20 ROIs, showed distinct clusters. Profiles were largely similar for all regions in the dMPFC subsystem (Figure 4.4B, upper left), while dMPFC itself was somewhat separated from the cluster, being displaced towards aMPFC. In addition, the activation profile for L-LTC resembled the MTL as well as the other regions in the dMPFC subsystems. Regions in the MTL network also had largely similar profiles (Figure 4.4B, middle), but with other notable features. vMPFC resembled not only other MTL regions, but also aMPFC, while for pIPL, there was high similarity not only to other MTL regions, but also to much of the dMPFC subsystem and conspicuously also to PCC. Within the core regions, aMPFC had a relatively distinct profile, but was most similar to other frontal regions, while PCC instead showed results closely similar to those of pIPL, with similarity to all other regions except for aMPFC, dMPFC, and TempP.

These results are summarized in the MDS plot in Figure 4.4C. As expected, regions of the dMPFC network largely cluster together, but with dMPFC shifted towards other frontal regions. Regions of the MTL network are again close together, with vMPFC somewhat apart from the main cluster. PCC, instead of clustering with its partner core region, is placed between dMPFC and MTL networks, in a position close to pIPL. aMPFC occupies a position between the other two frontal regions, as perhaps expected from anatomical proximity.

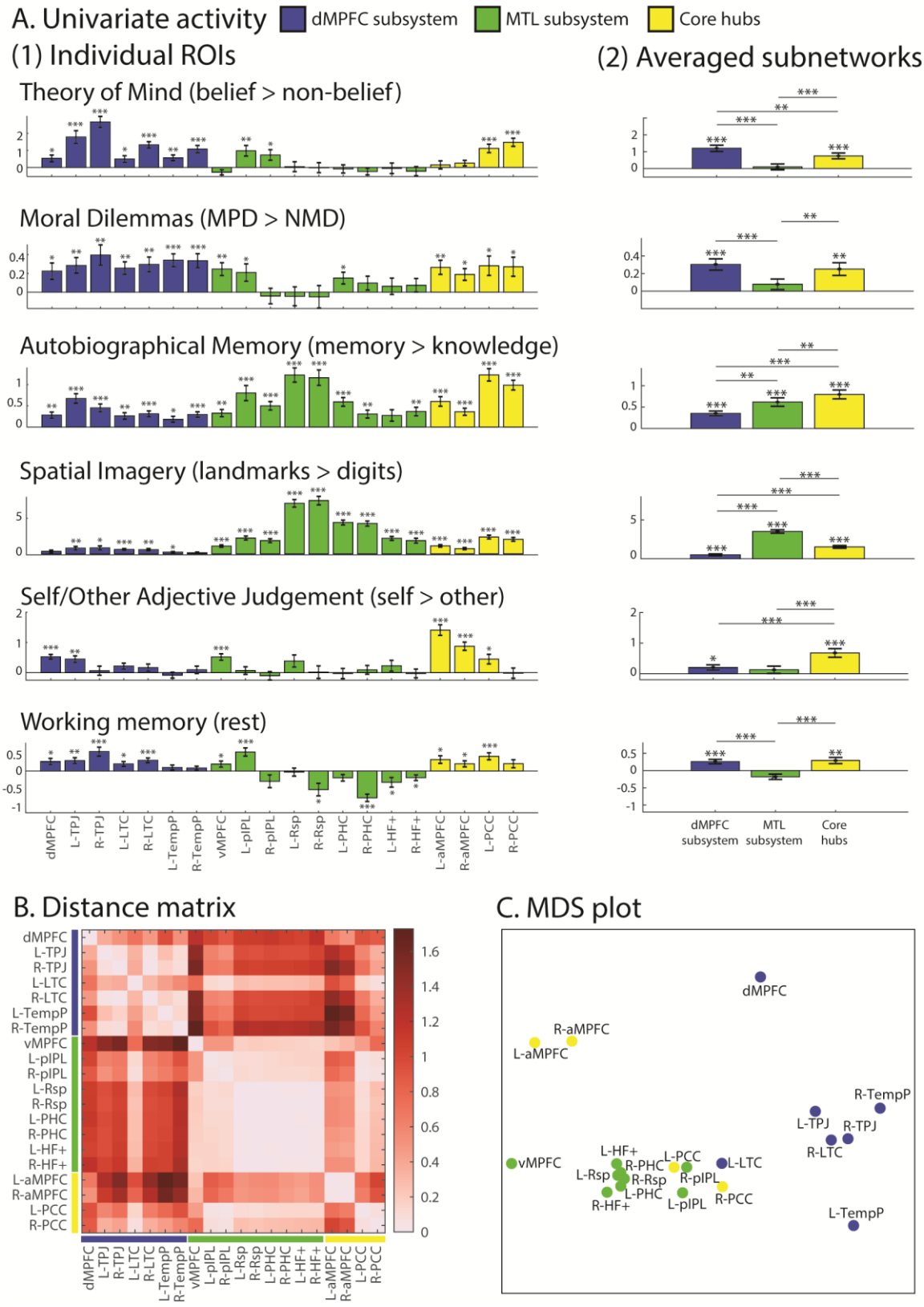


*Figure 4.4. (A) DMN ROIs recruited by each condition within the 6 tasks. Error bars represent standard error. t-tests against zero were conducted for each contrast in each (1) ROI or (2) subnetwork. *** indicates $p < 0.001$, ** indicates $p < 0.01$, and * indicates $p < 0.05$ (all tests were corrected for multiple comparisons using FDR). (B) Dissimilarity matrix calculated using 1 – Pearson's r between ROIs based*

on their activity profile across the 6 tasks. (C) Multidimensional scaling (MDS) to visualize the dissimilarity between regions.

4.4.4 Task-wise multi-voxel pattern similarity

To compare the similarity of voxelwise activity patterns across tasks (e.g., belief > non-belief vs. self > other), we correlated patterns of beta-values across voxels, for each pair of tasks, within each ROI and subnetwork (Figure 4.5A). Four model similarity matrices were constructed to test (1) whether the two "introspection of mental states" tasks were especially similar, (2) whether the two "memory-based construction/simulation" tasks were especially similar, (3) whether the self/other adjective judgment task was especially dissimilar to other contrasts, and (4) whether rest > working memory was especially dissimilar to other contrasts. Results showed that the dMPFC subsystem (dMPFC, R-TPL, R-LTC, and R-TempP), as well as pIPL and PCC had strong pattern similarity between the two "introspection" tasks. On the other hand, the MTL subsystem (pIPL, Rsp, PHC, HF+), as well as aMPFC and PCC showed strong pattern similarity between the two "memory-based construction" tasks. Across many ROIs of the three subsystems, there was a strong tendency for the self > other pattern to be distinct from others (greater similarity for contrast pairs not involving self/other). Few regions, however, showed the rest > working memory pattern to be distinct from the others (only R-Rsp). Together, these data complement the findings in Figure 4.4. Though regions in each subsystem contain voxels responding to each contrast, the pattern of these activations is organized along the lines proposed by Andrews-Hanna (2012), with more dissimilar activation patterns for contrasts predominantly associated with different networks.

The distance matrix (Figure 4.5B) and MDS plot (Figure 4.5C), based on correlations of the pattern-similarity matrices shown in Figure 4.5A, showed distinct clusters, largely similar to those based on univariate activity profiles. The ROIs of the dMPFC subsystem clustered with each other, as did many of the MTL ROIs. Again, however, PCC and IPL regions clustered close together, between dMPFC and MTL clusters, and again, despite putative assignment to different networks, there was some similarity of the three MPFC regions.

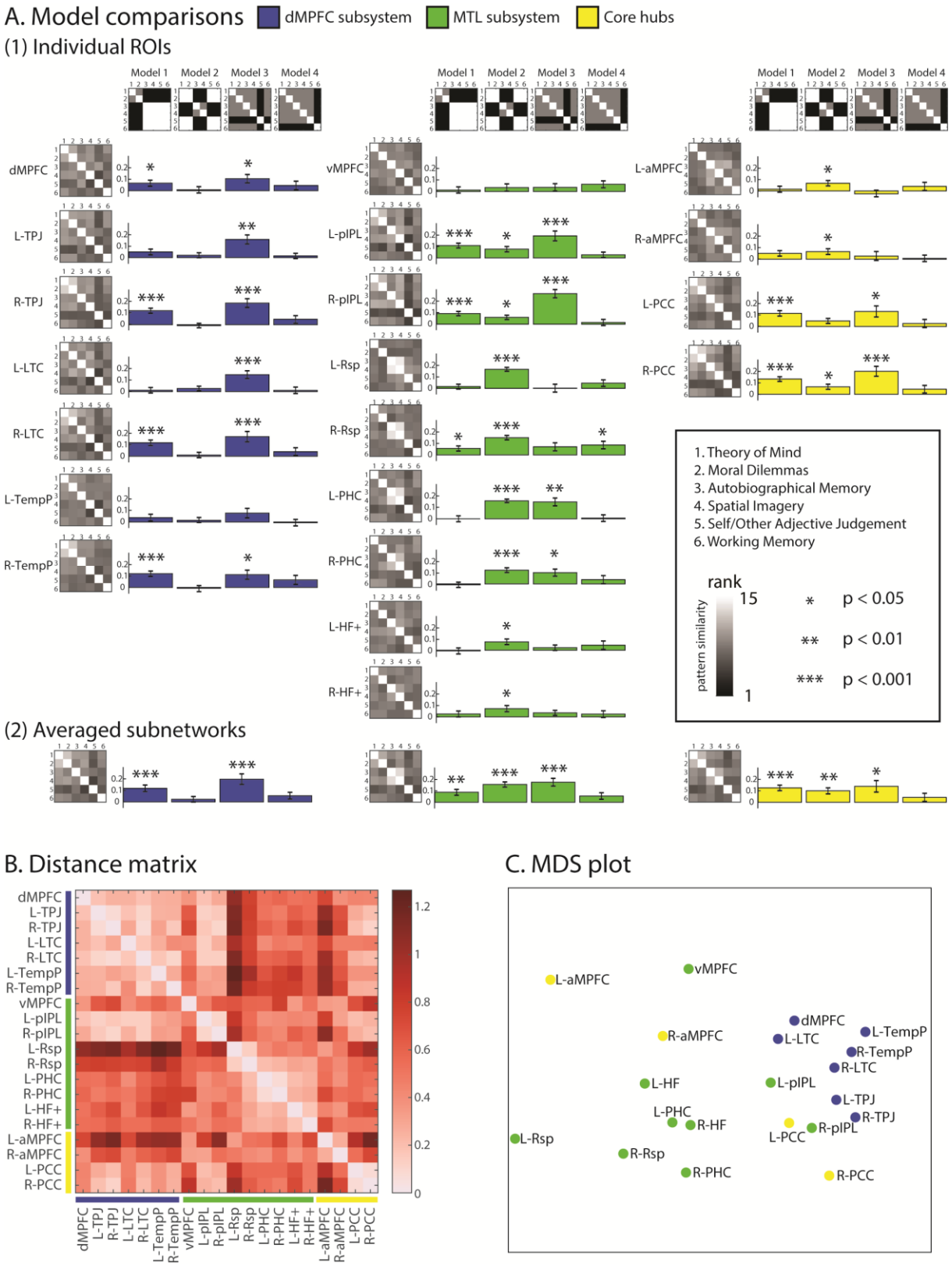


Figure 4.5. (A) Correlations between each pair of activation patterns in each (1) ROI (subnetworks in columns) or (2) subnetwork. The upper row shows the four model similarity matrices (white indicates empty cells that were not used in the comparisons, grays indicate 1s, and blank indicate 0s). In each matrix, tasks are ordered (top to bottom and left to right) as follows: theory of mind, moral dilemmas,

*autobiographical memory, spatial imagery, self/other and working memory. Leftmost columns show the average rank-transformed similarity matrices across subjects. Bar graphs represent Kendall's tau-a correlation between each participant's empirical and model similarity matrices tested against zero (corrected for multiple comparisons at FDR < 0.05). *** indicates p < 0.001, ** indicates p < 0.01, and * indicates p < 0.05. (B) Dissimilarity matrix calculated using 1 – Pearson's correlation between ROIs based on their correlation profiles across 15 task pairs. (C) Multidimensional scaling (MDS) to visualize the dissimilarity between regions.*

4.5 Discussion

Many complex cognitive processes have been linked to the DMN, supporting its role in high-level thought (Buckner and Carroll, 2007; Buckner et al., 2008; Spreng et al., 2009; Andrews-Hanna, 2012; Andrews-Hanna et al., 2014b). Among the most established of these cognitive functions are social, semantic, episodic, and self-relevant processing (Frith and Frith, 2006; Binder et al., 2009; McDermott et al., 2009; Spreng et al., 2009; Humphreys and Lambon Ralph, 2017). Recent findings suggest that the DMN consists of anatomically and functionally heterogeneous subsystems (Andrews-Hanna, 2012; Andrews-Hanna et al, 2014b; Yeo et al., 2011; Braga et al., 2017; Axelrod et al., 2017). Here, we used six diverse tasks to examine functional similarities and differences between DMN regions.

In many respects, our results matched the tripartite division proposed by Andrews-Hanna et al. (2010b, 2014b; 2012). In terms of univariate activity, regions of the dMPFC subsystem had largely similar activity profiles (Figure 4.4B), with strong response to our two social tasks, consistent with a particular role in social cognition or introspection about mental states. A partial exception was dMPFC itself, whose activity profile was shifted towards that of aMPFC (Figure 4.4B, C). In addition to their strong response to social contrasts, however, dMPFC regions also showed some response to most other contrasts (Figure 4.4A). Thus, specialization was quantitative rather than qualitative. Analysis of multivoxel activity patterns also largely supported the proposals of Andrews-Hanna et al. (2010b, 2014b; 2012), with regions of the dMPFC subsystem showing similar voxelwise activity patterns for our two social contrasts (Figure 4.5A), and again, largely similar profiles of between-task distances (Figure 4.5B, C).

Our results also support the proposal of an MTL subsystem, though with some caveats. In terms of univariate activity, regions of the MTL subsystem had largely similar activity profiles (Figure 4.4B), with strong response to the autobiographical memory and spatial

imagery tasks, and in most cases little response to other contrasts (Figure 4.4A). The most conspicuous exceptions were vMPFC, whose activity profile was shifted towards that of aMPFC, and pIPL, which responded to most contrasts (Figure 4.4B, C). Analysis of multivoxel patterns showed a largely similar picture. For MTL regions except vMPFC, voxelwise activity patterns were especially similar for the memory and imagery contrasts (Figure 4.5A), and across all task pairs, there were largely similar profiles of between-task distances (Figure 4.5B, C). Again, though, the distance profile of pIPL was rather different, with some similarity to other regions of both MTL and dMPFC subsystems, and again, vMPFC was shifted towards aMPFC (Figure 4.5B, C).

Our results give less support to the concept of a midline core consisting of aMPFC and PCC. In terms of both univariate and multivariate activity, aMPFC was more similar to the adjacent dMPFC and vMPFC regions than to PCC. In terms of univariate activity, its strongest response was to the self-other contrast (Figure 4.4A). In contrast, both univariate and multivariate analyses placed PCC between dMPFC and MTL subsystems, with results closely similar to those of pIPL (Figure 4.4C, 4.5C). If anything, these results suggest pIPL and PCC as a DMN functional "core" (consistent with Buckner et al. (2008, 2009)), while MPFC regions show some dorsal-ventral gradient but also resemblances to one another, and relatively distinct profiles compared to the other ROIs, including PCC.

Some important caveats should be considered. Undoubtedly, our *a priori* ROIs would not match the exact functional regions of individual participants, meaning that results for adjacent regions will to some extent blur together. One region where this consideration could be especially significant is the inferior parietal lobule, represented here by pIPL and TPJ ROIs (Figure 4.2). Our univariate data agreed with the proposals of Andrews-Hanna et al. (2010b) in broadly separating pIPL and TPJ. At a finer scale, however, it is possible that pIPL should be further subdivided, as suggested by some functional connectivity data (Yeo et al., 2011). Blurring of functionally separate regions within the pIPL might contribute to our findings of similarity to both dMPFC and MTL subsystems, resembling PCC. Similar considerations apply to our finding of broad similarities between the three MPFC regions. Of particular relevance here are the results of Braga & Buckner (2017), who scanned four individuals 24 times using fMRI. The authors found that two distinct networks, showing resemblance with the dMPFC and MTL subsystems in Andrews-Hanna et al. (2010), could be identified in each individual. However, a unique finding from this study was that spatially juxtaposed regions of the two networks were found in each of the three MPFC regions, dMPFC, aMPFC, and vMPFC, which

may be blurred together by spatial averaging in a group analysis. Despite these concerns, our results confirmed a dorsal-ventral gradient within the MPFC, with the dMPFC being more more involved in tasks requiring "introspection of mental states" and vMPFC more involved in tasks requiring "memory-based construction/simulation".

Other aspects of our results cannot be explained by spatial blurring. In particular, a conspicuous result was a significant response to non-social contrasts throughout most regions of the dMPFC subsystem, including those far from the MTL or core hubs. Along with the broad similarity of whole-brain maps for each contrast (Figure 4.3), apart from self > other, such results confirm partial, but not complete separation of response patterns for different DMN subsystems.

As noted earlier, several authors have proposed that the DMN represents broad features of a cognitive episode, situation or context (Hassabis and Maguire, 2007; Ranganath and Ritchey, 2012; Manning et al., 2014). Our results suggest both partial functional separation but also integration within this context representation. Matching many other findings (Andrews-Hanna et al., 2014a; Axelrod et al., 2017), our results link regions of the dMPFC subsystem to social cognition, and regions of the MTL subsystem to spatial or scene representation. To represent a cognitive episode, it is plausible that social and spatial representations are often integrated, for example to indicate who is where in the represented episode. Such integration may be achieved through communication between dMPFC and MTL subsystems, perhaps especially mediated by the pIPL and PCC. The self is also likely to be a core part of any episode representation, perhaps especially dependent on MPFC. In this way, the DMN acts partly as an integrated whole, but binding together aspects of the episode representation that are predominantly contributed by separate subregions.

Two other regions are worthy of further consideration. The first is the inferior frontal gyrus (IFG), which was not part of our *a priori* ROIs. Our whole-brain results (Figure 4.3A) showed that although IFG activity was weak in second-level analyses for most tasks (with the exception of self > other), a substantial minority of individual participants showed reliable recruitment for most tasks (Figure 4.3B). In the semantic literature, it has been shown that the semantic network, including the IFG, consists of many regions overlapping with the DMN (Binder et al., 2009; Noonan et al., 2013). In a dataset of 1000 participants, Yeo et al. (2011) identified the IFG as part of the dMPFC subsystem (Andrews-Hanna et al., 2014b). Given these findings, future studies should consider further the relation between the IFG and the DMN.

The second region requiring further consideration is the hippocampus. The hippocampal peak (HF+) defined in Andrews-Hanna et al. (2010b) is not located in the hippocampus proper, but lies between the PHC and perirhinal cortex (PRC) (Moore et al., 2014; Ritchey et al., 2015; https://neurovault.org/collections/3731/). The PHC has been linked to the "posterior medial system", a network closely related to the DMN, while the PRC has been linked to the "anterior temporal system", along with the temporal poles and orbitofrontal cortex (Ranganath and Ritchey, 2012). The role of the current HF+ ROI is therefore unclear as it may span functionally heterogeneous regions. Another question is whether the hippocampus is part of the DMN at all. Our results show a mixed picture, as only some contrasts activated parts of the hippocampus. Although the hippocampus has been associated with episodic memory and spatial navigation (Maguire et al., 1998; Addis et al., 2007; Rugg et al., 2012; Brown et al., 2016), it has been proposed to play a different role from other regions in the MTL subsystem. In particular, the hippocampus may integrate information across the anterior temporal and posterior medial systems (Ranganath and Ritchey, 2012).

Our findings provide a mixed answer to the question of functional specialization within the DMN. On the one hand, there is evidence of a largely integrated whole, with similar whole-brain activity maps for multiple contrasts, and some response to every contrast in each of the proposed subsystems, supporting classical accounts (e.g. Buckner and Carroll, 2007; Spreng et al., 2009). On the other hand, there is partial functional separation, in close accord with the proposals of separate dMPFC and MTL subsystems (Andrews-Hanna et al., 2010b, 2014a; Andrews-Hanna, 2012), though with remaining uncertainties over the concept of a midline core. Integrating social, spatial, self-related, and other aspects of a cognitive situation or episode, the DMN may provide the broad context for current mental activity.

Chapter 5 Discussion

The experiments described in this thesis explored representations in the human brain at varying levels of temporal granularity and abstraction. In Chapter 2, an experiment using high-temporal resolution EEG/MEG showed that different components of attentional representation unfold sequentially throughout an event within hundreds of milliseconds as a behaviorally-relevant target is selected. Chapter 3 employed fMRI to examine to what extent different brain regions represent individual items, steps, and episodes over a sequence of goal-directed actions. The visual cortex, MD network, and DMN network were all found to code for the currently relevant event, while the DMN additionally represented the broader episode. The DMN has been observed to be involved in many high-level abstract tasks, and Chapter 4 examined functional overlaps and specificity across different cognitive domains. The current studies, however, are insufficient to give a conclusive picture of the cognitive and neural mechanisms within the cortical networks that were investigated. This chapter will discuss various unexplored issues and possibilities arising from these studies.

5.1 Dynamic versus sustained coding of latent events

While the results in Chapter 2 showed multiple representational components during attentional selection of a visual target, the representation during the period between the cue onset and visual presentation did not seem to extend beyond the auditory representation of the cue. Yet it is during this period that participants set up a mental representation of the target. As discussed earlier in the context of preparatory attention, classical models of working memory have proposed that persistent spiking supports maintenance of working memory during a delay period, such that the neural firing patterns during the delay resemble the patterns in response to seeing the target itself (Fuster and Alexander, 1971; Chelazzi et al., 1993a). However, recent empirical data looking at single trial analysis of neural recordings in visual cortex (Kucewicz et al., 2017), parietal cortex (Shafi et al., 2007; Cromer et al., 2010), and prefrontal cortex (Shafi et al., 2007; Stokes et al., 2013; Lundqvist et al., 2016b) suggest that spiking activity during the delay period can be very sparse, and that activity travels through a continuous series of states rather than maintaining a stable pattern across time (Stokes et al., 2013; Stokes, 2015). Furthermore, while neuronal recordings have shown that delay activity is not locally persistent,

noninvasive EEG recordings of global activity have also revealed that, for extended periods of time, information held in working memory cannot be decoded (e.g., Wolff et al., 2015, 2017).

A statically sustained template may also be theoretically unlikely, as it has mechanistic problems, such as vulnerability to interference from distractors (Lundqvist et al., 2010, 2018; Miller et al., 2018). In laboratory experiments, the delay period is usually associated with a blank screen with a fixation cross, where there is little interference. However, under real-world conditions, we are constantly processing a continuous stream of incoming information; such that it would be difficult for neural patterns to maintain stability while processing new information at the same time. Parthasarathy et al. (2017) have shown that classifiers trained on times before an additional input do not generalize well to the times following it.

Another issue is that persistent spiking is metabolically expensive (Lundqvist et al., 2018; Miller et al., 2018). Although there are some examples in the literature showing single neurons that seem to show persistent activity on individual trials (e.g., Funahashi et al., 1989; Meyer et al., 2011), it has been shown with multiple-electrode studies that the majority of neurons spike sparsely (Lundqvist et al., 2018). The timescales used in working memory experiments are typically a few seconds at most. While it is conceivable for a neuron to persistently fire for short timescales, would we expect the neuron to persistently fire over several hours if the experiment demanded it?

Recently, there has been much theoretical development in how information is held in working memory (Stokes, 2015; Miller et al., 2018); however, it is still unknown how long-term memory and temporally extended episodes are coded. Imagine one morning you open the fridge and notice that the milk has run out and you tell yourself that you will need to buy milk on your way back from work in the evening. The requirement to buy milk will need to be stored somewhere in memory while you go about your day with many intervening activities and distracting information. The time between morning and leaving work may be several hours long. In our subjective experience, one would not actively maintain the mental representation of the goal in working memory throughout the entire period, but we are still able to maintain it for when it is needed. Is the coding of this long-term prospective memory mechanistically similar to how working memory is stored? Or does it involve a different mechanism? To investigate this, one would need to record high temporal resolution data over long timescales in naturalistic settings.

As reviewed in the Introduction, Baldassano et al. (2017) used naturalistic movie clips to examine how events are represented in the brain with fMRI. Another paper using the same approach has found similar results with EEG (Silva et al., 2019). Both studies suggest that events are represented in high-level regions as stable patterns of neural activity over an extended timescale, in spite of fluctuations in ongoing sensory input, only to rapidly shift at event boundaries. This is an interesting finding, which provides a neural mechanism of how slowly drifting thoughts, such as representations of situation models or schemas, are represented. However, as reviewed in https://nikokriegeskorte.org/2016/12/12/, although those papers claim that events are represented by stable patterns, it is a model assumption rather than a result demonstrated by the data. To explore to what extent events during naturalistic perception are coded using dynamic or stable patterns, cross-temporal generalizations of event representation using high temporal resolution recordings will be additionally needed.

5.2 How is information transferred across regions?

In Chapters 2 and 3, large scale cortical networks, including sensory, MD, and DMN regions are engaged during behaviorally complex tasks, but representing different kinds of information to different degrees and at different times. In particular, in Chapter 3, it was found that while sensory and MD regions are especially sensitive to the currently relevant information, DMN regions additionally represented the broader task episode. This finding suggests that as we experience the world, our thoughts are evolving at different timescales simultaneously, such that we are not thinking just one thing at a time (Manning et al., 2014; Baldassano et al., 2017; Manning, 2019). Most laboratory experiments have relied on trial-based paradigms, where participants are presented with independent displays of words or images, to examine how the brain responds to a single display in isolation. But to fully understand how the brain represents the world, it is insufficient to look at only one isolated moment; instead, we need to consider how that moment is temporally, semantically, and spatially related to other moments during the experience (Manning et al., 2014). If different brain regions use different timescales to together represent our experience, then these regions would need to have connections to transfer information to one another.

Studies of temporal processing hierarchies have led to the hypothesis that high-level regions accumulate and integrate information over time that is inputted into primary sensory regions (Hasson et al., 2008, 2015; Lerner et al., 2011). Baldassano et al. (2017) showed that

events are nested in a hierarchical structure from low level to high-level regions, as a significant proportion of boundaries in a given layer were also present in lower layers of the hierarchy. In an elegant study, Yeshurun et al. (2017) found that the brain accumulates and integrates locally small word changes to construct unique neural representations for different stories. They constructed distinct narratives by changing only a few words in each sentence (e.g., "he" to "she"; "sobbing" to "laughing"), and found that neural responses to small word changes become increasingly amplified in long-timescale regions. These studies provide evidence that high-level regions have access to information from lower level regions, but additional studies are needed to investigate how information is transferred, and to what extent low-level regions can access information from high-level regions.

A few recurrent neural network models have been built to explore what mechanisms might support nested hierarchical structure of timescales. For example, Chaudhuri et al. (2015) constructed a large-scale dynamical model based on anatomical connectivity in the macaque neocortex. By varying the density of excitatory connection strengths for each area according to the position of that area in the cortical hierarchy, their model was able to show that decay times in response to stimulus input increased progressively from sensory areas to association areas in the hierarchy (Chaudhuri et al., 2015; Chen et al., 2015). Chung et al. (2017) developed a novel updating mechanism in a recurrent neural network, where chunks of information are transmitted from lower to higher levels primarily at event boundaries. Their model was able to capture temporal dependencies by discovering the latent hierarchical structure of the inputted sequence.

In empirical data, one approach to examining information flow is effective (or directed) connectivity (Friston et al., 1993; Friston, 2011). A commonly used approach is Granger causality (Granger, 1988; Ding et al., 2006). According to Granger causality, time series X Granger-causes time series Y if the past values of X help predict the future of Y better than the past values of Y alone (Seth et al., 2015; Bastos and Schoffelen, 2016). One can therefore test whether time series in one region predicts the activation in another region at a later time. Goddard et al. (2016) extended this idea to examine how one region's representational patterns influence another region's representational patterns across time. The authors identified an early feedforward flow and later feedback between peri-frontal and peri-occipital sensors during object recognition. Other methods, such as dynamic causal modeling, can also examine directed connectivity between multiple brain regions, and comparing a large range of possible models, finally selecting the best model using Bayesian model comparison (Friston et al., 2003). It would be interesting to explore using such methods how different regions within the cortical

processing hierarchy transfer information over extended timescales. A tentative Granger causality analysis was performed using the data from Chapter 2, but encountered difficulties defining the optimal size of time windows and number of previous time points to put in the model. Similar challenges must be overcome when examining information flow across temporal hierarchies.

It has been shown that functional connectivity patterns in large-scale networks can flexibly reassemble to meet a variety of task demands (Cole et al., 2013a). More recent research has shown that activations evoked by cognitive tasks can be predicted in held-out brain regions via estimated activity flow over resting-state functional connectivity networks (Cole et al., 2016). This suggests that resting-state functional connectivity networks shape activity flow in task contexts. Building on this finding, the authors developed a new technique, known as information transfer mapping, to quantify the amount of information transferred between pairs of brain regions via resting-state functional connectivity. This approach accurately predicted the activation pattern in response to a task in a target region based on a source region's activation pattern (Ito et al., 2017). However, the task they used was a laboratory task which required participants to respond to three rule domains (logic, sensory, and motor) on a trial-by-trial basis. It is unknown whether this approach would be able to predict information flow across cortical networks for continuous naturalistic data where information is unfolding at multiple timescales.

5.3 Progression through steps, and special status of the first and the last step

Although not explored in detail in the experimental chapters, there could be multiple ongoing processes that evolve as one progresses through a task episode. Similar to previous studies (e.g., Farooqui et al., 2012; Desrochers et al., 2015) that observed an increase in activity in various regions throughout the task episode, the experiment in Chapter 3 showed an increasing linear trend in MD and DMN regions using FIR analysis. However, after separately characterizing onset and epoch responses, different networks showed different linear trends of increasing and decreasing onset and epoch activity. It is difficult to completely interpret how these univariate changes relate to cognition, and the various progressive cognitive processes during execution of a task episode may not be mutually exclusive. It is unknown to what extent such processes have contributed to the results in Chapter 3, as well as to other studies. More targeted experimental designs would be necessary to distinguish these processes, and future

studies would benefit from designs that directly examine changes in information representation over the course of an episode. A few possible processes will be discussed here.

Decreasing cognitive load over time

Anderson & Matessa (1990) described a production system theory of serial memory, suggesting that serial lists are represented as hierarchical structures in which items are embedded, and that recall depends on a limited-capacity activation process (also see Baddeley, 1990) where activation is divided among elements within the hierarchical structure. Their empirical data from a task where participants memorized lists of digits that were chunked during presentation, showed that recall latency depended on chunking of the items and increased according to the length of the sequence (Anderson and Matessa, 1990). The idea of a hierarchical control structure can be extended to active task sequence execution. Such studies have shown that reaction times are longer when executing events marking the beginning of a task sequence, compared to identical events within the sequence, and the reaction time increase for the first step is directly related to the complexity of the sequence (Rogers and Monsell, 1995; Wylie and Allport, 2000; Schneider and Logan, 2006; Farooqui and Manly, 2018b). For example, Schneider & Logan (2006) found that the sequence initiation time for the sequence ABBA is longer than AABB (where A and B represent task rules), where although the two sequences contained the same elements they differed in rule transition frequency.

Based on these observations, Farooqui & Manly (2018b) suggested that representations of task episodes are assembled at the beginning of execution and are initially represented as one cognitive unit. They further suggested that, consequently, the cognitive load related to the mental program is highest at the beginning of the episode and decreases as more parts of the episode are executed (Farooqui and Manly, 2018a, 2018b). Supporting their hypothesis, they found that deactivation in the DMN (a region known to deactivate in response to cognitive load) was strongest at the start of the episode, but gradually returned to baseline as sequential steps were completed. Similar decrease in deactivation was also found in the DMN epoch regressors in Chapter 3; and it was additionally observed that the onset regressor was strongest at the beginning of the first step (possibly related to retrieval) and decreased over time.

Increasing cognitive control over time

Desrochers et al. (2015) found that the rostrolateral prefrontal cortex exhibited a ramping pattern of activation as a function of progression through a task sequence that was not influenced by task complexity or switching. To examine the role of the increased rostrolateral

prefrontal activity, the authors performed a follow-up experiment where TMS was delivered at different temporal positions within the task sequence. Results showed that rostrolateral frontal stimulation resulted in increased error rates that increased as a function of position. Combining these findings, the authors proposed that the increasing rostrolateral prefrontal activity over time reflects a source of top-down control to resolve uncertainty throughout the sequence. If task representations are assembled or refreshed at the initiation of the episode, there is no uncertainty of position at the first event; however, as each subsequent step is executed, there is a greater probability of losing track of one's position. Therefore, greater cognitive control may be needed at further steps into the episode.

Farooqui et al. (2012) proposed another possible explanation for why additional cognitive control is required as the episode progresses. Using a hierarchically structured task episode paradigm where the episode contained subtasks (e.g., search for "CAT" then "X"), they observed in MD regions that activity was greater when completing a higher order task. The authors suggested that a task episode is characterized by a unique neurocognitive configuration of multiple mental processes. Therefore, completion of each task would result in a revision of the mental program, and the completion of a higher level task would require a greater revision.

End of task episode

In Chapter 3, the DMN regions showed rapid increase of activity at the end of the task episode, which can be seen from the FIR timecourses, and characterized by the offset regressor. This finding is consistent with studies of event segmentation, which showed increased DMN activity at time points where participants perceived an event boundary (Speer et al., 2007; Ben-Yakov et al., 2013; Baldassano et al., 2017; Ben-Yakov and Henson, 2018). It has been suggested that the posterior medial regions of the DMN are involved in coding for situation models, and a rapid shift in the contents of the situation would result in changes in DMN activity (Baldassano et al., 2017; DuBrow et al., 2017; Brunec et al., 2018). It is possible that different regions within the DMN increase their activity at event boundaries for different reasons, such as reflecting prediction error or encoding of the event into memory (Baldassano et al., 2017; Brunec et al., 2018; Silva et al., 2019). Future studies would be needed to disentangle the various possibilities.

5.4 Dissociating action sets and task goals

Chapter 3 showed that regions within the DMN code for task episodes. However, a limitation of this study is that each episode contained unique sets of actions to achieve completion; therefore, we could not know whether different brain regions differently code for sequences of items and task goals. For example, when making coffee, one might need to boil water, add sugar, and add milk. However, this collection of acts is not restricted to the goal of making coffee. Other schemas, for example making tea, could require carrying out the parts of the same sequence of actions. This example illustrates that task episodes are more than the sum of each individual act, and our mental programs must be able to represent the current schema in order to carry out the correct goal. Furthermore, it is possible for different action sets to lead to the same goal. For example, if the goal is to make coffee, one could make instant coffee, French press coffee, or go to Starbucks to buy coffee. It is possible that some brain regions could represent the general goal regardless of the specific actions or number of steps it took to reach the goal.

Further experiments will be needed to understand how the brain represents task goals (e.g., making coffee vs. making tea) and action sets (e.g., boil water, add sugar, and add milk). A possible way to approach this would be to build on the experiment in Chapter 3. The proposed experiment is presented in Figure 5.1, where the same items (from step 2 and onwards) can occur in different goals, and the same goal can be achieved by assembling different items. While the task structure would be similar to the experiment in Chapter 3, a prompt would be needed at the end of the task episode to verify that participants had kept the correct goal in mind. Additionally, it would be interesting to vary the number of steps required to complete a task episode, such that goal completion can be distinguished from step number (e.g., step 3 in sequences "add instant coffee → add water → drink coffee" and "add instant coffee → add water → add milk → drink coffee" differ in whether the goal is completed).

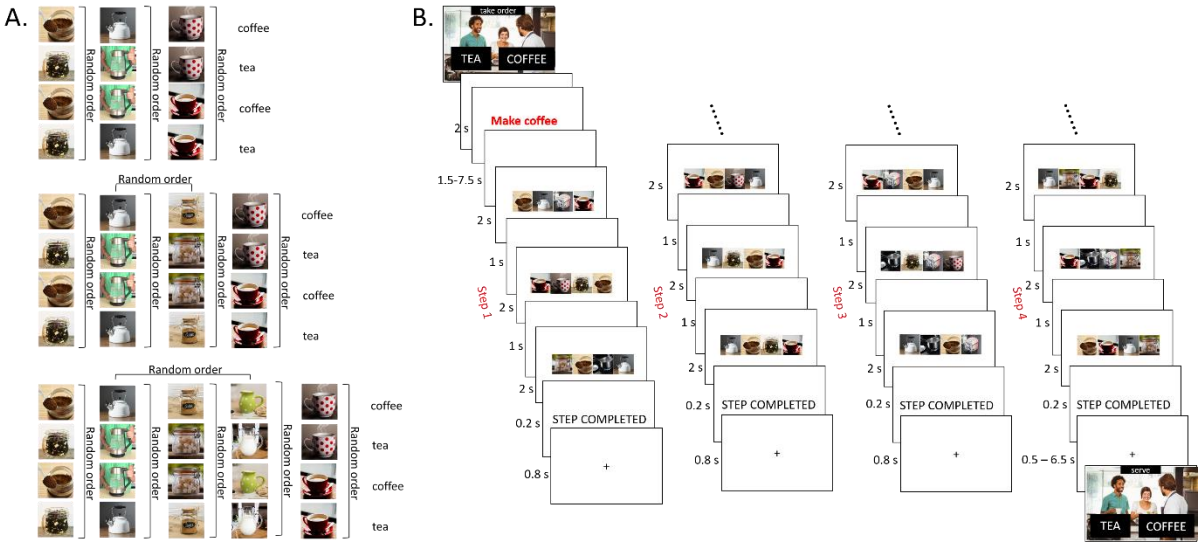


Figure 5.1 Extension of the experiment in Chapter 3 to dissociate action sets and task goals. (A) stimulus sets used in the experiment. Same schemas can contain different items, and the same sequences of items (from step 2 and onwards) can belong to different schemas. (B) Structure of an example task episode. The participant would be asked to choose at the end of the episode the correct schema to validate they had kept the scheme in mind.

It is a plausible hypothesis that the posterior medial network and MPFC are differently sensitive to action sets and task goals. The posterior medial network is proposed to be involved in representing situation models (Ranganath and Ritchey, 2012). A situation model is a higher-level cognitive representation of relationships between different elements of a particular episode. Schemas differ from situation models in that they describe extracted commonalities across multiple episodes and lack unit detail (Radvansky and Zacks, 2011; Ghosh and Gilboa, 2014b). It has been proposed that the MPFC accumulates information about the context of interrelated episodes, and is involved in the representation of schemas (Preston and Eichenbaum, 2013; Ghosh and Gilboa, 2014b; Spalding et al., 2015; Gilboa and Marlatte, 2017; Robin and Moscovitch, 2017). Studies have suggested a dissociation between the MPFC and MTL when encoding new information, such that MPFC acts to detect the congruency of new information with existing schemas, while MTL captures novel experiences and incongruent information (van Kesteren et al., 2012, 2013). For example, patient studies found that those with focal lesions within the MTL show congruency benefits when learning new information that is consistent with existing schema, while patients with additional lesions in semantic areas do not show such benefits (Kan et al., 2009). A dissociation between these regions was shown in a recent fMRI study (Reagh, SfN 2018), where participants were shown movie clips of actors

in two different cafes and two different grocery stores. The researchers found that the posterior medial network showed increased pattern similarity for repetitions of movies of the same context compared to related and different contexts, and the MPFC showed increased pattern similarity for both same and related contexts but not different contexts. These observations suggest that the posterior medial network may code for specific detailed episodes, while the MPFC codes for more general schemas. Future research will be needed to test whether this dissociation occurs in active goal-directed tasks. In Chapter 3, there was a hint of room encoding in MPFC, but it did not survive FDR correction. However, the tasks in the same room were quite different from each other, and all had different goals. It is plausible that MPFC coding for room would increase if there were a clearer abstraction that differentiated the goals between kitchen and bathroom tasks.

5.5 What are the functions of the DMN?

In Chapter 3, the focus was on the role of the DMN in representing task episodes. However, in Chapter 4, it was shown that the DMN was engaged during the performance of multiple tasks from various cognitive domains (including theory of mind, solving moral dilemmas, autobiographical memory, spatial navigation, and judging self attributes). The DMN was also found to be more active during a passive rest versus a working memory task. Due to its involvement in vastly different tasks as well as passive rest, it is difficult to pinpoint its engagement to a specific function. The role of the DMN in human cognition is still debated, and some of the proposed hypotheses will be reviewed below.

The DMN has often been characterized as a network that increases its activity during passive rest compared to externally-oriented, goal-directed tasks (Shulman et al., 1997; Mazoyer et al., 2001; Raichle et al., 2001). It has also been found to show increased activity during easy tasks compared to difficult tasks (Mazoyer et al., 2002; McKiernan et al., 2003, 2006). Such studies suggest that the DMN is associated with mental processes that occur during rest but are suspended during cognitive tasks (Raichle et al., 2001; Christoff et al., 2004; Goldberg et al., 2006). Two possibilities are illustrated in Figure 5.2. As described by William James (1890) as the “stream of consciousness”, when left without an immediate task that demands full attention, our minds wander from one passing thought to the next (Figure 5.2A). Using thought sampling and fMRI, researchers have found that people’s minds tend to wander during rest or when task demands are low, and individuals’ reports of the tendency of their

minds to wander correlated with activity in the DMN (Christoff et al., 2004; McKiernan et al., 2006; Mason et al., 2007). The context of mental activity during rest has often been reported to be associated with reminiscence of past experiences, making future plans, and other personal thoughts and experiences (Andreasen et al., 1995; Binder et al., 1999; Mazoyer et al., 2001; Andrews-Hanna et al., 2010a). These observations show that the DMN is active during spontaneous introspective cognition, also referred to as stimulus-independent thought or mind-wandering (Buckner et al., 2008; Andrews-Hanna, 2012).

However, while the DMN is appreciated for its elevated activity during passive control states, most of these paradigms consist of tasks that require performing cognitive operations on external sensory input on the current trial. Reviews of the literature reveal that experimental tasks that require participants to engage in various forms of internal mentation also engage the DMN (Buckner and Carroll, 2007; Buckner et al., 2008; Andrews-Hanna, 2012; Andrews-Hanna et al., 2014b). A quantitative meta-analysis focused on four domains, including autobiographical memory, navigation, theory of mind, and "default mode" showed relatively high correspondence with the DMN (Spreng et al., 2009). This was followed by a direct test of some of these domains in a single experimental setup (Spreng and Grady, 2010), again revealing a common neural pattern resembling the DMN. Although autobiographical memory, prospection, navigation, and theory of mind seem to be very different cognitive domains, Buckner & Carroll (2007) observed that each requires the ability to mentally project oneself from the present moment into a simulation of another time, place, or perspective. Goal-directed tasks that engage DMN activity also share in common the process of one's attention being inward, as opposed to external stimuli (Golland et al., 2008; Spreng et al., 2010; Andrews-Hanna, 2012). These observations are consistent with the "internal mentation hypothesis" (Buckner et al., 2008; Andrews-Hanna, 2012) that proposes that the DMN contributes directly to internal mentation, in particular of self-relevant events (Kelley et al., 2002; Buckner et al., 2008; Benoit et al., 2010).

Another hypothesis is the "sentinel hypothesis" which suggests that the DMN functions to support monitoring of internal and external environments when focused attention is relaxed (Figure 5.2B; Shulman et al., 1997; Gusnard and Raichle, 2001; Gusnard et al., 2001; Gilbert et al., 2006, 2007; Buckner et al., 2008). Studies of the PCC indicate a role in continuously gathering information about the world around us including sensory events, and also one's behavior in the service of spatial orientation and memory (Vogt et al., 1992; Gusnard and Raichle, 2001). Most experimental tasks require focused attention on the imperative stimulus,

and by contrast, passive conditions release the participant to broadly monitor the environment. This broad low-level focus has been termed an "exploratory state" (Shulman et al., 1997) or "watchfulness" (Gilbert et al., 2007). Several studies have found that increased DMN activity during baseline was predictive of faster reaction times in the upcoming trial in diffuse attention conditions, suggesting a passive role in monitoring the environment (Small et al., 2003; Gilbert et al., 2006; Hahn et al., 2007).

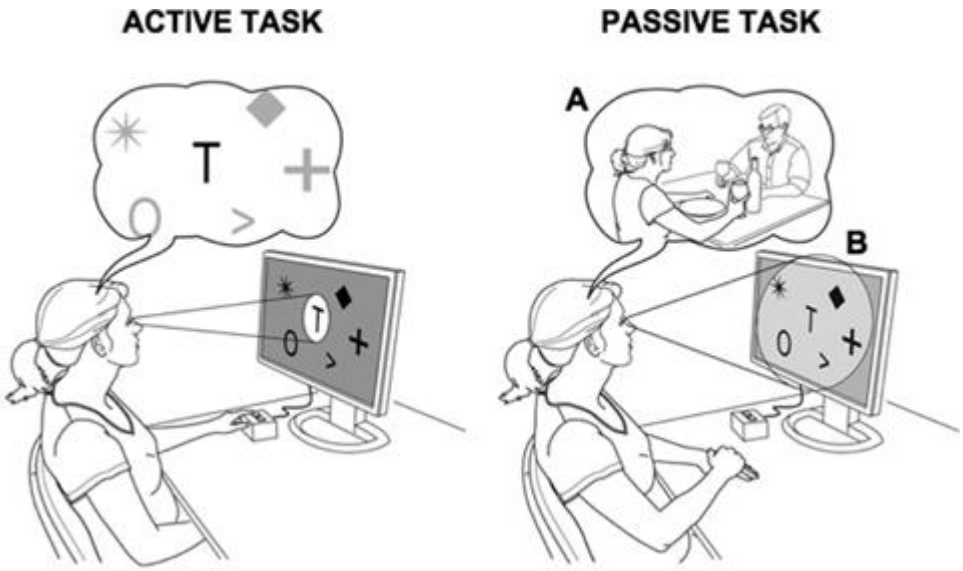


Figure 5.2 The functions of the default network have been difficult to unravel because passive tasks, which engage the default network, differ from active tasks on multiple dimensions. As one goes from an active task demanding focused attention (left panel) to a passive task (right panel), there is both a change in mental content (A) and level of attention to the external world (B). Spontaneous thoughts unrelated to the external world increase (A). There is also a shift from focused attention to a diffuse low-level of attention (B). Hypotheses about the functions of the default network have variably focused on one or the other of these two distinct correlates of internally directed cognition. Reproduced from Buckner et al. (2008).

The "internal mentation hypothesis" and "sentinel hypothesis" may not be mutually exclusive, and some research emphasizes the importance of representations of both the self and the environment (Stawarczyk et al., 2011; Mantini and Vanduffel, 2013). Both our internal and external awareness may reflect complementary and coexisting aspects of conscious experience (Mantini and Vanduffel, 2013). It has been suggested that the DMN plays a role in supporting conscious experience, and it has been shown that altered states of consciousness, such as

anesthesia (Martuzzi et al., 2010), sleep (Sämann et al., 2011), and out-of-body experiences (De Ridder et al., 2007) result in changes to DMN activity and connectivity.

A similar yet alternative view of summarizing the DMN's function in representing both internal mentation and external environment is that it represents context (Manning et al., 2014). Researchers have argued that the posterior medial network, including the PHC, and other regions of the DMN, is involved in contextual representation (Diana et al., 2007; Ranganath, 2010b; Ranganath and Ritchey, 2012). However, context is one of the most fundamental yet elusive concepts in memory research, as researchers often define context by exclusion: in an experiment, there is a set of *items* that the participant is asked to memorize, and there is *context* which reflects everything else that is represented in the participant's brain during the experiment (Manning et al., 2014). Context, therefore, may include, for example, information about the external environment, mood, thoughts about recently encountered events, contents of mind-wandering, and incidental features of the encoded stimuli (Smith and Vela, 2001; Manning et al., 2014). Under this definition, anything could be context. For example, Wang et al. (2013) examined functional dissociations in the PHC and PRC by experimentally manipulating meaningless visual fractal stimulus pairs, where one fractal in each pair was treated as an item and the other as context during encoding. The context fractals were shown in the background and remained on the screen for four consecutive trials, while the item fractals were transiently presented in the foreground. In accordance with the BIC model, they found that the PHC was involved in representing context and PRC in representing item, despite the context and item sharing the same stimulus material. These studies suggest that "context" is a useful way to summarize the contents of representations within the DMN; however, a better operational definition for context would be helpful to systematically study it. One possible definition of context is as slowly varying information, a perspective which is discussed in the next section.

5.6 Temporal context

Temporal context has been a core theme in this thesis, and will be elaborated here. Temporal context can be defined according to the timescale of information processing, and can be operationally defined as slowly drifting mental representation (Manning et al., 2014). It has been suggested that generally speaking, there are two different (and non-mutually exclusive) ways for mental context to slowly drift (Manning et al., 2014). The first could arise from the

brain representing slowly drifting features of the external world, for example, as one travels through space (because we cannot physically teleport, our location in one moment will be similar to our location in the next moment, and similarly for the objects we observe). The second source is our internal mentation, which the common phase "train of thought" well characterizes the interconnection in the sequence of thought where one idea leads to another. As reviewed in the introduction, studies using audiovisual movies have shown a hierarchy of temporal receptive windows in the human cortex, with the shortest receptive windows in the sensory regions, and long timescale processing corresponding to high-level regions, in particular the DMN (Hasson et al., 2008; Lerner et al., 2011).

According to this view of temporal context (Howard and Kahana, 2002; Polyn and Kahana, 2008; Manning et al., 2014), slowly-drifting contextual information serves to organize more transient or quickly changing information (items). For example, when recalling an event, people tend to not only recover the features of the event itself, but they also recover the information associated with other events that occurred nearby in time. Manning et al. (2011) showed that the events and thoughts surrounding a target event may be considered a context for the target event. The authors isolated candidate context representations by identifying temporally autocorrelated neural patterns recorded from ECoG as patients studied a list of words. During recall of an item, the authors found that neural patterns matched not only the target item itself, but showed similarity to neighboring items, with similarity decreasing as a function of temporal distance. Similar results were found in a related study, where extracellular recordings were taken from MTL regions of epileptic patients (Howard et al., 2012). These findings provide neural match to the contiguity effects in episodic memory, and align with retrieved context models suggesting that contiguity effects are a result of a gradually changing state of temporal context (Howard and Kahana, 2002; Polyn and Kahana, 2008; Howard and Eichenbaum, 2013).

Furthermore, it has been suggested that MTL, DMN, and some PFC regions play a crucial role in representing the sequences of events that are experienced (Jenkins and Ranganath, 2010; Howard and Eichenbaum, 2013; Hsieh et al., 2014; Cohn-Sheehy and Ranganath, 2017). Although our everyday experiences often overlap in the contents of events (e.g., turning on the tap to wash vegetables and turning on the tap to wash your face), representation of a sequence is one way to address the differentiation between these overlapping events. It has been shown that MTL and DMN regions show increased activity when successfully learning and retrieving overlapping sequences (Kumaran and Maguire, 2006; Ross

et al., 2009; Brown et al., 2010). With MVPA, Hsieh et al. (2014) found that the hippocampal activity patterns are more similar for adjacent items in a learned sequence, but hippocampal patterns for objects were not affected by temporal proximity in a random sequence. In the same study, the authors also found that the hippocampal activation patterns are sensitive to sequence boundaries, showing decreased pattern similarity between items when transitioning to another learned sequence. These findings are in line with those suggesting DMN regions exhibit long temporal receptive windows (Hasson et al., 2008; Lerner et al., 2011), and exhibit stable patterns of activity within an event (Baldassano et al., 2017). These studies suggest that the DMN is involved in a higher-order conceptual representation of events, which transcends the immediate environment, and can bind information across multiple overlapping events (Murphy et al., 2018).

It is plausible that temporal context is just one form of context that is represented in the DMN. Although the concept of context may be difficult to define, some insights could be derived from previous studies. In the Context Maintenance and Retrieval (CMR) model (Polyn et al., 2008), the authors describe context as "a pattern of activity in the cognitive system, separate from the pattern immediately evoked by the perception of a studied item, that changes over time and is associated with other coactive patterns". Examples of context could be source (e.g., Johnson, 2006; Mitchell and Johnson, 2009), task (e.g., Polyn et al., 2009; Crittenden et al., 2015), semantic (e.g., Polyn et al., 2005), or temporal information (e.g., Ezzyat and Davachi, 2014; Hsieh et al., 2014; Manning et al., 2016). These studies have shown that regions of the DMN, in particular, MTL and PFC regions are involved in context representation. fMRI studies have shown that a shift in task or temporal context is associated with increased DMN activity (Speer et al., 2007; Crittenden et al., 2015; Baldassano et al., 2017). Furthermore, it has been shown that the semantic network highly overlaps with the DMN, especially with the core hubs and the dMPFC subsystem (Binder et al., 2009; Humphreys et al., 2015), and that similarity of neural patterns of pictured objects or words in these areas reflected their semantic similarity (Fairhall and Caramazza, 2013). A characteristic of mental context is that it groups individual items in terms of similarity/association along a particular dimension (e.g., semantic context may group "nurse, ambulance, medicine" and "flower, shovel, hose" as hospital and gardening contexts; temporal context may group together events with close temporal proximity in the same context). These observations fall in line with Badre et al.'s (Badre and D'Esposito, 2007; Badre, 2008; Badre and Nee, 2018) ideas of representational and temporal abstraction, with representations higher in the hierarchy recruiting higher-level regions. More research would

be needed to establish the overlapping and different neural mechanisms for representing temporal, task, and semantic contexts.

5.7 Implications and extensions

This thesis has explored representational dynamics within a single event and across temporally extended episodes during purposeful behavior. It has also examined possible functional roles and the types of representations that occur within sensory regions, the MD network, and the DMN. The experiments conducted in this thesis share interesting overlaps with other cognitive domains, and a few of these overlaps will be discussed below.

Biased competition in other domains – memory as an example

The experiment in Chapter 2 examined the timecourses of component processes of selective attention. When attending for a target, visual input must compete for cognitive resources, and features most relevant to the target gain the most attentional weight. When multiple objects compete, an enhanced representation of a particular object will be at the expense of other objects' representations (Duncan, 1996, 2006; Beck and Kastner, 2009). This is known as the biased competition theory of selective attention. For example, Siedl et al. (Seidl et al., 2012) investigated with fMRI how such target category enhancement and distractor suppression is implemented at a neural level during visual search. They found using pattern similarity analysis that target information was significantly increased, while the information about the distractor object was significantly decreased, relative to a neutral category. However, competition is a quite ubiquitous phenomenon, and can occur in other cognitive domains when mental representations are competing for limited resources.

Competition is found in memory when a cue is linked to more than one item in memory; in this situation, those items compete for access to conscious awareness (Anderson et al., 1994; Anderson and Neely, 1996). Any negative effect on memory performance associated with competition is referred to as interference. For example, competition can be experienced in daily life when you are prompted to type in a password on the computer and accidentally type an old password even though you have changed it already to a new one. Anderson and colleagues' hypothesis suggests that successfully retrieving a target item depends not only on how strongly the cue is associated with the target, but also on whether the cue is related to other items in memory as well (Anderson and Neely, 1996). During competition, inhibitory control mechanisms are engaged to suppress the distractor items during selective memory retrieval, and

can lead to the forgetting of competing memories (Anderson et al., 1994; Anderson and Spellman, 1995; Levy and Anderson, 2002). Several paired-associate paradigms, such as the AB/AC paradigm have been used to examine memory competition. In these experiments, participants are given cue-target pairs to remember, such as "fruit-orange", and later another pair with a shared cue, such a "fruit-apple". Researchers found that selectively practicing retrieval on one pair causes impaired recall of the other pair compared to a control pair that was never practiced (Anderson et al., 1994). Using multivariate pattern analysis in fMRI, researchers have shown simultaneous reactivation of competing memories during selective retrieval, and that greater reactivation of the competing memory was associated with less accurate target memory (Kuhl et al., 2011, 2012). Furthermore, Wimber et al. (2015) developed a canonical template tracking method to quantify the activation state of individual target and competitor memories during repeated retrieval practice, and found that repeatedly retrieving target memories lead to enhanced representation of target patterns and suppression of competitor patterns over time.

The consequence of memory competition is enhanced memory for the target compared to distractors, but it is assumed that multiple component processes are involved. For example, behavioral studies have found that there are some precise conditions necessary for distractor inhibition to occur, including target retrieval, competitor interference, and recall testing (Anderson and Spellman, 1995). It is particularly interesting to note that competitor interference is necessary for distractor inhibition to occur. Studies have found that distractor inhibition requires a moderate similarity of targets and competitors, and there would be facilitation for competitors sharing highly overlapping features, while very little competition for highly distinct competitors (Anderson and Spellman, 1995; Anderson et al., 2000). This phenomenon has been characterized by a U-shaped function relating the degree of memory activation to change in memory strength. One account is known as the nonmonotonic plasticity hypothesis, where compared to lower and higher levels of memory activation, moderate levels of activation lead to weakening of the memory (Detre et al., 2013). A prediction of the nonmonotonic plasticity hypothesis is that close competition can result in strengthening of the winning memory and trigger weakening of memories that lose the competition. This has been shown using MVPA analysis in fMRI, where trials that elicited similar levels of classifier evidence (indicating close competition) for two items held in working memory resulted in worse subsequent memory performance (Lewis-Peacock and Norman, 2014).

It has been proposed that biased competition in memory retrieval and selective attention share the same mechanisms (Anderson and Spellman, 1995). However, research on neural mechanisms of memory competition has so far mostly used fMRI. It would be interesting to characterize the timecourse of memory competition with high temporal resolution decoding methods, and in particular to test the hypothesized serial effect of competitor interference followed by enhancement/inhibition of target and distractor, as well as test which phases of competition are related to subsequent memory.

Navigation

The cognitive map hypothesis, originally proposed by Tolman (1948), describes how the brain forms a systematically organized representation of the environment that can be used to flexibly guide purposeful behavior. Research on cognitive maps originally highlighted the roles of the hippocampus, MTL, and posterior medial regions in navigation (e.g., Hartley et al., 2003; Spiers and Maguire, 2006; Spiers and Gilbert, 2015). However, more recent research has demonstrated that the mechanisms involved in navigation may apply to navigating in nonphysical spaces, and that the cognitive maps may be applied broadly to many cognitive domains (Schiller et al., 2015; Epstein et al., 2017), including social (Tavares et al., 2015), conceptual (Constantinescu et al., 2016), and temporal space (Ezzyat and Davachi, 2014; Ekstrom and Ranganath, 2017).

When carrying out a task episode, one must keep the goal in mind (e.g., "make a stew") while performing the immediate tasks at hand (e.g., "take food from fridge", "wash vegetables", etc.) as one navigates towards the goal. Real world navigation can be seen as cognitively similar to performing a task with sequential steps. Similarly, one would need to keep a destinational goal in mind (e.g., "London") while traveling the routes to get there (e.g., "walk to train station", "sit on the train", etc) and interacting with various objects along the way.

Relevant to the experiment described in Chapter 3, a spatial experiment has been conducted by Kim & Maguire (2018). Participants in the study were passively moved in a 3D virtual building, and the authors found that the anterior hippocampus represented local information within a room, and the RSC, PHC, and posterior hippocampus represented room information within the wider building. Their findings of DMN regions showing a greater representation of the wider building compared to local information is analogous to the DMN in representing episodes compared to items. Another experiment by Brown et al. (2016) employed virtual reality to investigate how the brain codes prospective navigational goals. Participants

learned five goal locations before the scan, and inside the scanner, participants began each trial at one of the locations and were cued to plan navigation of the shortest route to the cued goal before actively navigating. MVPA during the planning period showed that the hippocampus coded for both current location and future goals, as well as for intervening locations along the route. Furthermore, the strength of hippocampal goal representations covaried with goal-related coding in MPFC, MTL, and RSC. These findings illuminate the mechanistic role of the hippocampus, MTL, and DMN regions in hierarchical coding for environments as well as goal-directed planning.

Most of the past research on navigation has focused on relationships between object and space during passive viewing or planning, and more research would be needed to investigate how the brain codes for information during active navigation. For example, the MTL is important in not only navigating via familiar routes but also in enabling the person to come up with new shortcuts through path integration (McNaughton et al., 2006; Howett et al., 2019; e.g., if a person moves from location 1 to location 2 to location 3 that are positioned in a triangular configuration, it is possible to figure out how to go from location 3 to location 1 in a straight line even though the person has never traveled that route). One interesting follow-up would be to use a virtual reality paradigm to look at how the hippocampus and interrelated cortical structures support prospective representation of navigational goals, routes, and orientation when planning a shortcut or traveling through the original route. It would be possible to ask questions such as which brain regions code for the goal regardless of the route, and which brain regions code for overlapping routes regardless of the goal? Another interesting extension would be to examine how the brain solves detours in navigation (Spiers and Gilbert, 2015) and tasks (for example, when a learned route is blocked, or an ingredient has run out when making a stew, and additional actions are required to reach the goal). Detours are especially interesting, as they often first take the person away from the goal in order to be able to reach the goal subsequently. It could enable us to ask questions such as which brain regions keep the goal in mind, and which brain regions code for (absolute versus stepwise) distance towards the goal?

Real-world tasks

A topic that has been lightly touched on in previous sections is laboratory versus real-world naturalistic experiments. Laboratory experiments often require participants to make decisions on shapes or words on a trial-by-trial basis; however, in real-world settings, our ongoing sensory experience and internal thoughts are continuous and embedded within each other. In real life tasks, such as “making a stew” or “baking cupcakes”, one must be able to self-

initiate switches between various sub-activities while keeping the overall rules and goals in mind (Burgess, 2000; Logie et al., 2010; Rothbart and Posner, 2015). Several key studies have found that patients with rostral prefrontal lobe damage (in approximately Brodmann area 10) show deficits in performing everyday tasks (such as planning a dinner or doing grocery shopping), even though some had superior IQ and intact performances on neuropsychological tests of attention, memory, and executive functions (also known as 'dysexecutive syndrome'; Shallice and Burgess, 1991; Burgess, 2000; Roca et al., 2011). This suggests read-world tasks require an aspect of executive function that is not captured by most standardized laboratory tests (Manly et al., 2002; Roca et al., 2009, 2011, 2012).

Several tests have been designed to emulate everyday goal management demands, including the Six Elements Test (Shallice and Burgess, 1991), the Multiple Errands Test (Shallice and Burgess, 1991), and the Hotel Test (Manly et al., 2002). In the Hotel Test, for example, participants were asked to imagine working in a hotel and were asked to perform six subtasks within 15 minutes, which included compiling individual bills, sorting the charity collection, looking up telephone numbers, sorting conference labels, proofreading the hotel leaflet, and opening and closing garage doors. As completing all tasks is designed to take much longer than 15 minutes, participants were asked to do something from each within the total time available and spend as long as possible on each of the six tasks. The critical measure of the test is the management of the higher-level goal (to complete all tasks) when simultaneously dealing with competing demands of ongoing tasks.

However, not many neuroimaging studies have used naturalistic active tasks, possibly due to previous analytical challenges (Spiers and Maguire, 2007) and obvious practical constraints given that these tasks often involve participants actively moving to manipulate objects or travel to various locations (Logie et al., 2011). A few computer-based tasks have been developed to explore realistic tasks in the laboratory, including the Virtual Errands Test (McGeorge et al., 2001), and the Edinburgh Virtual Errands Test (Logie et al., 2011), and, recently, Cullen et al. (2016) created a Computerised Multiple Elements Test to investigate goal management in the scanner. In a pilot study, Cullen et al. (2016) contrasted brain activity when participants were required to manage switching across all tasks in the Computerised Multiple Elements Test compared to a condition where switching was prompted, and found that the rostral prefrontal cortex was more active in the former condition.

Many more interesting questions can be explored by bringing naturalistic tasks into the scanner; indeed, many key neuroscience questions cannot be addressed effectively in

conventional laboratory paradigms (Maguire, 2012). For example, Manly et al. (2002) found that by giving patients interrupting tones during the Hotel Task, patients were able to perform better on the task. The authors suggested that perhaps the tones drew attention away from the task at hand and the higher-level goal could be expressed. This reinstatement of the latent goal is especially interesting. Previous studies have shown that working memory can be "activity silent" (Stokes, 2015), but if a perturbation is administered (by giving participants a high-contrast task-irrelevant stimulus during the maintenance period), the contents of working memory become decodable (Wolff et al., 2015, 2017). It would be interesting to test whether interrupting tones would similarly make high-level goals more decodable. Another interesting research avenue would be to apply event models such as the model developed in Baldassano et al. (2017) to explore which brain regions are involved in managing different tasks as well as which regions are involved in self-generated compared to experimenter-prompted switches.

Temporal perception and clinical implications

The thesis has investigated some of the functional roles of the MD and DMN networks. Menon (2011) suggested a unifying triple network model of psychopathology, including in this model the frontoparietal central executive network (FPN; which is part of the MD network, and includes MFG, dlPFC, and IPS), DMN, and salience network (SN; also another part of the MD network and includes AI and ACC). Among the many networks identified by resting-state analyses, these networks have been particularly important in understanding cognitive function. Menon's (2011) review suggested that deficits in access, engagement, or disengagement of these three networks play a prominent role in multiple neurological disorders, including schizophrenia, depression, anxiety, bipolar, dementia, autism, etc. However, most studies have mainly focused on group differences within regions or individual networks, and more studies are needed to understand the relationships between networks, and how they relate to behavior in clinical populations.

The studies reviewed, as well as the experiment in Chapter 3, suggest a distinction between the MD and DMN networks in their sensitivity to fine details and coarse context. Although everyday life consists of a stream of continuous information, people are able to experience and selectively retrieve discrete events in memory, suggesting that event segmentation is an integral part of how we experience the world (Zacks et al., 2001b, 2007; Reynolds et al., 2007). The understanding of human cognition, and how it can go wrong, would benefit from studying temporal episodes, beyond trial-by-trial studies. In an interesting applied study, Zacks et al. (2006) examined the relationship between event perception and memory in

Alzheimer patients. Compared with young adults and healthy age-matched controls, patients with Alzheimer's disease showed deficits in parsing ongoing activity into appropriate events, in particular, poor agreement with others' event segmentation was associated with poorer memory for stimuli used in the tasks, as well as poor memory for the temporal order of the events. It would be interesting to extend this type of research into fMRI to understand the involvement of these functional networks in event perception in other populations, which may give insight into how they experience the world.

5.8 Concluding remarks

This thesis explored human cognition at multiple scales. On one hand, the chapters characterized brain regions involved in increasingly abstract levels of cognitive thought, ranging from simple visual selection to executing task episodes to complex social cognition. On the other hand, there had been much focus on processing timescales. EEG/MEG and fMRI have been used to characterize dynamic representational changes over time. Although the physical world is continuously changing, and our inner thoughts drift in and out, we are able to integrate as well as segment this information into meaningful units. The cortical organization of both processing timescales and cognitive abstraction seem to form a processing hierarchy from sensory regions to the MD network and finally to the DMN. As we interact with the world, these regions must work together to together represent our "experience".

References

Addis DR, Pan L, Vu M-A, Laiser N, Schacter DL (2009) Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. Neuropsychologia 47:2222–2238.

Addis DR, Wong AT, Schacter DL (2007) Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. Neuropsychologia 45:1363–1377.

Anderson JR, Matessa M (1990) A Production System Theory of Serial Memory.

Anderson MC, Bjork RA, Bjotk E (1994) Remembering Can Cause Forgetting: Retrieval Dynamics in Long-Term Memory Article. J Exp Psychol Learn Mem Cogn.

Anderson MC, Green C, Mcculloch KC (2000) Similarity and Inhibition in Long-Term Memory: Evidence for a Two-Factor Theory. 26:1141–1159.

Anderson MC, Neely JH (1996) Interference and Inhibition in Memory Retrieval. Memory:237–313.

Anderson MC, Spellman BA (1995) On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. Psychol Rev 102:68–100.

Andreasen NC, O'Leary DS, Cizadlo T, Arndt S, Rezai K, Watkins GL, Ponto LL, Hichwa RD (1995) Remembering the past: two facets of episodic memory explored with positron emission tomography. Am J Psychiatry 152:1576–1585.

Andrews-Hanna JR (2012) The brain's default network and its adaptive role in internal mentation. Neuroscientist 18:251–270.

Andrews-Hanna JR, Reidler JS, Huang C, Buckner RL (2010a) Evidence for the Default Network's Role in Spontaneous Cognition. J Neurophysiol 104:322–335.

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL (2010b) Functional-Anatomic Fractionation of the Brain's Default Network. Neuron 65:550–562.

Andrews-Hanna JR, Saxe R, Yarkoni T (2014a) Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses. Neuroimage 91:324–335.

Andrews-Hanna JR, Smallwood J, Spreng RN (2014b) The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. Ann N Y Acad Sci 1316:29–52.

Asaad WF, Rainer G, Miller EK (2000) Task-Specific Neural Activity in the Primate Prefrontal

Cortex. J Neurophysiol 84:451–459.

Assem M, Blank I, Mineroff Z, Ademoglu A, Fedorenko E (2017) Neural Activity in the Fronto-Parietal Multiple Demand Network Robustly Predicts Individual Differences In Working Memory And Fluid Intelligence. bioRxiv:110270.

Awh E, Belopolsky A V., Theeuwes J (2012) Top-down versus bottom-up attentional control: a failed theoretical dichotomy. Trends Cogn Sci 16:437–443.

Awh E, Jonides J (2001) Overlapping mechanisms of attention and spatial working memory. Trends Cogn Sci 5:119–126.

Axelrod V, Rees G, Bar M (2017) The default network and the combination of cognitive processes that mediate self-generated thought. Nat Hum Behav 1:896–910.

Baddeley AD (1990) Human memory : theory and practice. Psychology Press.

Baddeley AD, Hitch G (1974) Working Memory. Psychol Learn Motiv 8:47–89.

Badre D (2008) Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. Trends Cogn Sci 12:193–200.

Badre D, D'Esposito M (2007) Functional Magnetic Resonance Imaging Evidence for a Hierarchical Organization of the Prefrontal Cortex. J Cogn Neurosci Early Acce:080219115128817–080219115128818.

Badre D, D'Esposito M (2009) Is the rostro-caudal axis of the frontal lobe hierarchical? Nat Rev Neurosci 10:659–669.

Badre D, Nee DE (2018) Frontal Cortex and the Hierarchical Control of Behavior. Trends Cogn Sci 22:170–188.

Baldassano AC, Hasson U, Norman KA (2018) Representation of real-world event schemas during narrative perception.

Baldassano C, Chen J, Zadbood A, Pillow JW, Hasson U, Norman KA (2017) Discovering Event Structure in Continuous Narrative Perception and Memory. Neuron 95:709-721.e5.

Baldauf D, Desimone R (2014) Neural Mechanisms of Object-Based Attention. Science (80-) 344:424–427.

Bar M (2007) The proactive brain: using analogies and associations to generate predictions. Trends Cogn Sci 11:280–289.

Bartlett FC (Frederic C, Kintsch W (1932) Remembering : a study in experimental and social psychology. Cambridge University Press.

Bastos AM, Schoffelen J-M (2016) A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. Front Syst Neurosci 9:1–23.

Battistoni E, Stein T, Peelen M V. (2017) Preparatory attention in visual cortex. Ann N Y Acad

Sci:1–16.

Beck DM, Kastner S (2009) Top-down and bottom-up mechanisms in biasing competition in the human brain. Vision Res 49:1154–1165.

Ben-Yakov A, Dudai Y (2011) Constructing realistic engrams: poststimulus activity of hippocampus and dorsal striatum predicts subsequent episodic memory. J Neurosci 31:9032–9042.

Ben-Yakov A, Eshel N, Dudai Y (2013) Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. J Exp Psychol Gen 142:1255–1263.

Ben-Yakov A, Henson RN (2018) The Hippocampal Film Editor: Sensitivity and Specificity to Event Boundaries in Continuous Experience. J Neurosci 38:10057–10068.

Ben-Yakov A, Rubinson M, Dudai Y (2014) Shifting gears in hippocampus: temporal dissociation between familiarity and novelty signatures in a single event. J Neurosci 34:12973–12981.

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29:1165–1188.

Benoit RG, Gilbert SJ, Volle E, Burgess PW (2010) When I think about me and simulate you: Medial rostral prefrontal cortex and self-referential processes. Neuroimage 50:1340–1349.

Binder JR, Desai RH, Graves WW, Conant LL (2009) Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. Cereb Cortex 19:2767–2796.

Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW (1999) Conceptual Processing during the Conscious Resting State: A Functional MRI Study. J Cogn Neurosci 11:80–93.

Bor D, Duncan J, Wiseman RJ, Owen AM (2003) Encoding strategies dissociate prefrontal activity from working memory demand. Neuron 37:361–367.

Braga RM, Buckner RL (2017) Parallel Interdigitated Distributed Networks within the Individual Estimated by Intrinsic Functional Connectivity. Neuron 95:457-471.e5.

Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M (2008) Top-Down Control of Human Visual Cortex by Frontal and Parietal Cortex in Anticipatory Visual Spatial Attention. J Neurosci 28:10056–10061.

Brown TI, Carr VA, LaRocque KF, Favila SE, Gordon AM, Bowles B, Bailenson JN, Wagner AD (2016) Prospective representation of navigational goals in the human hippocampus. Science (80-) 352:1323–1326.

Brown TI, Ross RS, Keller JB, Hasselmo ME, Stern CE (2010) Which way was I going?

Contextual retrieval supports the disambiguation of well learned overlapping navigational routes. J Neurosci 30:7414–7422.

Brunec IK, Moscovitch M, Barense MD (2018) Boundaries Shape Cognitive Representations of Spaces and Events. Trends Cogn Sci 22:637–650.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: Anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38.

Buckner RL, Carroll DC (2007) Self-projection and the brain. Trends Cogn Sci 11:49–57.

Buckner RL, Sepulcre J, Talukdar T, Krienen FM, Liu H, Hedden T, Andrews-Hanna JR, Sperling RA, Johnson KA (2009) Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. J Neurosci 29:1860–1873.

Bundesen C (1990) A theory of visual attention. Psychol Rev 97:523–547.

Bundesen C, Habekost T, Kyllingsbæk S (2005) A Neural Theory of Visual Attention: Bridging Cognition and Neurophysiology. Psychol Rev 112:291–328.

Bundesen C, Vangkilde S, Petersen A (2015) Recent developments in a computational theory of visual attention (TVA). Vision Res 116:210–218.

Burgess PW (2000) Strategy application disorder: the role of the frontal lobes in human multitasking. Psychol Res 63:279–288.

Cabeza R, St Jacques P (2007) Functional neuroimaging of autobiographical memory. Trends Cogn Sci 11:219–227.

Callicott JH, Mattay VS, Bertolino A, Finn K, Coppola R, Frank JA, Goldberg TE, Weinberger DR (1999) Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. Cereb Cortex 9:20–26.

Carlson T, Tovar DA, Alink A, Kriegeskorte N (2013) Representational dynamics of object vision: The first 1000 ms. J Vis 13:1–1.

Chaudhuri R, Knoblauch K, Gariel M-A, Kennedy H, Wang X-J (2015) A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. Neuron 88:419–431.

Chawla D, Rees G, Friston KJ (1999) The physiological basis of attentional modulation in extrastriate visual areas. Nat Neurosci 2:671–676.

Chelazzi L, Duncan J, Miller EK, Desimone R (1998) Responses of Neurons in Inferior Temporal Cortex During Memory-Guided Visual Search. J Neurophysiol 80:2918–2940.

Chelazzi L, Miller EK, Duncan J, Desimone R (1993a) A neural basis for visual search in inferior temporal cortex. Nature 363:345–347.

Chelazzi L, Miller EK, Duncan J, Desimone R (1993b) A neural basis for visual search in inferior temporal cortex. Nature 363:345–347.

Chen J, Hasson U, Honey CJ (2015) Processing Timescales as an Organizing Principle for Primate Cortex. Neuron 88:244–246.

Christoff K, Gordon AM, Smallwood J, Smith R, Schooler JW (2009) Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. Proc Natl Acad Sci U S A 106:8719–8724.

Christoff K, Ream JM, Gabrieli JDE (2004) Neural Basis of Spontaneous thought Processes. Cortex 40:623–630.

Chung J, Ahn S, Bengio Y (2017) Hierarchical Multiscale Recurrent Neural Networks.

Cohn-Sheehy BI, Ranganath C (2017) Time regained: how the human brain constructs memory for time. Curr Opin Behav Sci 17:169–177.

Cole MW, Bagic A, Kass R, Schneider W (2010) Prefrontal dynamics underlying rapid instructed task learning reverse with practice. J Neurosci 30:14245–14254.

Cole MW, Etzel JA, Zacks JM, Schneider W, Braver TS (2011) Rapid Transfer of Abstract Rules to Novel Contexts in Human Lateral Prefrontal Cortex. Front Hum Neurosci 5:142.

Cole MW, Ito T, Bassett DS, Schultz DH (2016) Activity flow over resting-state networks shapes cognitive task activations. Nat Neurosci 19:1718–1726.

Cole MW, Laurent P, Stocco A (2013a) Rapid instructed task learning: A new window into the human brain's unique capacity for flexible cognitive control. Cogn Affect Behav Neurosci 13:1–22.

Cole MW, Reynolds JR, Power JD, Repovs G, Anticevic A, Braver TS (2013b) Multi-task connectivity reveals flexible hubs for adaptive task control. Nat Neurosci 16:1348–1355.

Cole MW, Yarkoni T, Repovs G, Anticevic A, Braver TS (2012) Global connectivity of prefrontal cortex predicts cognitive control and intelligence. J Neurosci 32:8988–8999.

Constantinescu AO, O'Reilly JX, Behrens TEJ (2016) Organizing conceptual knowledge in humans with a gridlike code. Science (80-) 352:1464–1468.

Cooper R, Shallice T (2000) Contention scheduling and the control of routine activities. Cogn Neuropsychol 17:297–338.

Corbetta M, Shulman GL (2002) Control of Goal-Directed and Stimulus-Driven Attention in the Brain. Nat Rev Neurosci 3:215–229.

Crittenden BM, Duncan J (2014) Task difficulty manipulation reveals multiple demand activity but no frontal lobe hierarchy. Cereb Cortex 24:532–540.

Crittenden BM, Mitchell DJ, Duncan J (2015) Recruitment of the default mode network during

a demanding act of executive control. Elife 2015:1–12.

Crittenden BM, Mitchell DJ, Duncan J (2016) Task Encoding across the Multiple Demand Cortex Is Consistent with a Frontoparietal and Cingulo-Opercular Dual Networks Distinction. J Neurosci 36:6147–6155.

Cromer JA, Roy JE, Miller EK (2010) Representation of Multiple, Independent Categories in the Primate Prefrontal Cortex. Neuron 66:796–807.

Cullen B, Brennan D, Manly T, Evans JJ (2016) Towards validation of a new computerised test of goal neglect: Preliminary evidence from clinical and neuroimaging pilot studies. PLoS One 11:1–12.

Cusack R, Mitchell DJ, Duncan J (2010) Discrete Object Representation, Attention Switching, and Task Difficulty in the Parietal Lobe. J Cogn Neurosci 22:32–47.

Cusack R, Vicente-Grabovetsky A, Mitchell DJ, Wild CJ, Auer T, Linke AC, Peelle JE (2015) Automatic analysis (aa): efficient neuroimaging workflows and parallel processing using Matlab and XML. Front Neuroinform 8:90.

D'Esposito M (2007) From cognitive to neural models of working memory. Philos Trans R Soc B Biol Sci 362:761–772.

Dale AM (1999) Optimal experimental design for event-related fMRI. Hum Brain Mapp 8:109–114.

De Ridder D, Van Laere K, Dupont P, Menovsky T, Van de Heyning P (2007) Visualizing Out-of-Body Experience in the Brain. N Engl J Med 357:1829–1833.

Dehaene S, Changeux JP (2011) Experimental and Theoretical Approaches to Conscious Processing. Neuron 70:200–227.

Desimone R, Duncan J (1995) Neural Mechanisms of Selective Visual. Annu Rev Neurosci 18:193–222.

Desrochers TM, Burk DC, Badre D, Sheinberg DL (2016) The Monitoring and Control of Task Sequences in Human and Non-Human Primates. Front Syst Neurosci 9:1–18.

Desrochers TM, Chatham CH, Badre D (2015) The Necessity of Rostrolateral Prefrontal Cortex for Higher-Level Sequential Behavior. Neuron 87:1357–1368.

Desrochers TM, Collins AG, Badre D (2018) Sequential control underlies robust ramping dynamics in the rostrolateral prefrontal cortex. J Neurosci:1060–18.

Detre GJ, Natarajan A, Gershman SJ, Norman KA (2013) Moderate levels of activation lead to forgetting in the think/no-think paradigm. Neuropsychologia 51:2371–2388.

Diana RA, Yonelinas AP, Ranganath C (2007) Imaging recollection and familiarity in the medial temporal lobe: a three-component model. Trends Cogn Sci 11:379–386.

Ding M, Chen Y, Bressler SL (2006) Granger Causality: Basic Theory and Application to Neuroscience.

Dixon ML, Fox KCR, Christoff K (2014) Evidence for rostro-caudal functional organization in multiple brain areas related to goal-directed behavior. Brain Res 1572:26–39.

Dixon ML, Girn M, Christoff K (2017) Hierarchical Organization of Frontoparietal Control Networks Underlying Goal-Directed Behavior. In: The Prefrontal Cortex as an Executive, Emotional, and Social Brain, pp 133–148. Tokyo: Springer Japan.

Dodell-Feder D, Koster-Hale J, Bedny M, Saxe R (2011) fMRI item analysis in a theory of mind task. Neuroimage 55:705–712.

Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE (2007) Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci 104:11073–11078.

Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE (2006) A Core System for the Implementation of Task Sets. Neuron 50:799–812.

Downar J, Crawley AP, Mikulis DJ, Davis KD (2000) A multimodal cortical network for the detection of changes in the sensory environment. Nat Neurosci 3:277–283.

Dubis JW, Siegel JS, Neta M, Visscher KM, Petersen SE (2016) Tasks Driven by Perceptual Information Do Not Recruit Sustained BOLD Activity in Cingulo-Opercular Regions. Cereb Cortex 26:192–201.

DuBrow S, Rouhani N, Niv Y, Norman KA (2017) Does mental context drift or shift? Curr Opin Behav Sci 17:141–146.

Dumontheil I, Thompson R, Duncan J (2011) Assembly and Use of New Task Rules in Fronto-parietal Cortex. J Cogn Neurosci 23:168–182.

Duncan J (1980) The locus of interference in the perception of simultaneous stimuli. Psychol Rev 87:272–300.

Duncan J (1996) Cooperating brain systems in selective perception and action (Inui T., McClelland JL, eds). Cambridge, MA, US: The MIT Press.

Duncan J (2001) An adaptive coding model of neural function in prefrontal cortex. Nat Rev Neurosci 2:820–829.

Duncan J (2006) EPS mid-career award 2004: Brain mechanisms of attention. Q J Exp Psychol 59:2–27.

Duncan J (2010) The multiple-demand (MD) system of the primate brain: mental programs for

intelligent behaviour. Trends Cogn Sci 14:172–179.

Duncan J (2013) The Structure of Cognition: Attentional Episodes in Mind and Brain. Neuron 80:35–50.

Duncan J, Chylinski D, Mitchell DJ, Bhandari A (2017) Complexity and compositionality in fluid intelligence. Proc Natl Acad Sci 114.

Duncan J, Humphreys G, Ward R (1997) Competitive brain activity in visual attention. Curr Opin Neurobiol 7:255–261.

Duncan J, Humphreys GW (1989) Visual search and stimulus similarity. Psychol Rev 96:433–458.

Duncan J, Owen AM (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci 23:475–483.

Eichenbaum H (2013) Memory on time. Trends Cogn Sci 17:81–88.

Eimer M (2015) EPS Mid-Career Award 2014: The control of attention in visual search: Cognitive and neural mechanisms. Q J Exp Psychol 68:2437–2463.

Eimer M, Grubert A (2014) The gradual emergence of spatially selective target processing in visual search: From feature-specific to object-based attentional control. J Exp Psychol Hum Percept Perform 40:1819–1831.

Ekstrom AD, Ranganath C (2017) Space, time, and episodic memory: The hippocampus is all over the cognitive map. Hippocampus:1–8.

Epstein RA, Patai EZ, Julian JB, Spiers HJ (2017) The cognitive map in humans: spatial navigation and beyond. Nat Neurosci 20:1504–1513.

Erez Y, Duncan J (2015) Discrimination of Visual Categories Based on Behavioral Relevance in Widespread Regions of Frontoparietal Cortex. J Neurosci 35:12383–12393.

Esterman M, Yantis S (2010) Perceptual expectation evokes category-selective cortical activity. Cereb Cortex 20:1245–1253.

Everling S, Tinsley CJ, Gaffan D, Duncan J (2002) Filtering of neural signals by focused attention in the monkey prefrontal cortex. Nat Neurosci 5:671–676.

Ezzyat Y, Davachi L (2011) What Constitutes an Episode in Episodic Memory? Psychol Sci 22:243–252.

Ezzyat Y, Davachi L (2014) Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. Neuron 81:1179–1189.

Fahrenfort JJ, Grubert A, Olivers CNL, Eimer M (2017) Multivariate EEG analyses support high-resolution tracking of feature-based attentional selection. Sci Rep 7:1886.

Fairhall SL, Caramazza A (2013) Brain regions that represent amodal conceptual knowledge. J Neurosci 33:10552–10558.

Farooqui AA, Manly T (2018a) Hierarchical Cognition Causes Task-Related Deactivations but Not Just in Default Mode Regions. eneuro 5:ENEURO.0008-18.2018.

Farooqui AA, Manly T (2018b) We do as we construe: extended behavior construed as one task is executed as one cognitive entity. Psychol Res:1–20.

Farooqui AA, Mitchell D, Thompson R, Duncan J (2012) Hierarchical Organization of Cognition Reflected in Distributed Frontoparietal Activity. J Neurosci 32:17373–17381.

Fedorenko E, Duncan J, Kanwisher N (2013) Broad domain generality in focal regions of frontal and parietal cortex. Proc Natl Acad Sci 110:16616–16621.

Finoia P, Mitchell DJ, Hauk O, Beste C, Pizzella V, Duncan J (2015) Concurrent brain responses to separate auditory and visual targets. J Neurophysiol 114:1239–1247.

Fletcher PC, Happé F, Frith U, Baker SC, Dolan RJ, Frackowiak RSJ, Frith CD (1995) Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition 57:109–128.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A 102:9673–9678.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorical Representation of Visual Stimuli in the Primate Prefrontal Cortex. Science (80-) 291:312–316.

Friston KJ (2011) Functional and Effective Connectivity: A Review. Brain Connect 1:13–36.

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1993) Functional Connectivity: The Principal-Component Analysis of Large (PET) Data Sets. J Cereb Blood Flow Metab 13:5–14.

Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. Neuroimage 19:1273–1302.

Frith CD, Frith U (2006) The Neural Basis of Mentalizing. Neuron 50:531–534.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 61:331–349.

Fuster JM, Alexander GE (1971) Excitation and inhibition of neuronal firing in visual cortex by reticular stimulation. Science 173:652–654.

Garner KG, Dux PE (2015) Training conquers multitasking costs by dividing task representations in the frontoparietal-subcortical system. Proc Natl Acad Sci 112:14372–14377.

Ghosh VE, Gilboa A (2014a) What is a memory schema? A historical perspective on current neuroscience literature. Neuropsychologia 53:104–114.

Ghosh VE, Gilboa A (2014b) What is a memory schema? A historical perspective on current neuroscience literature. Neuropsychologia 53:104–114.

Giesbrecht B, Weissman DH, Woldorff MG, Mangun GR (2006) Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. Brain Res 1080:63–72.

Gilbert SJ, Dumontheil I, Simons JS, Frith CD, Burgess PW (2007) Comment on "Wandering minds: the default network and stimulus-independent thought". Science (80-) 317:43; author reply 43.

Gilbert SJ, Simons JS, Frith CD, Burgess PW (2006) Performance-related activity in medial rostral prefrontal cortex (area 10) during low-demand tasks. J Exp Psychol Hum Percept Perform 32:45–58.

Gilboa A, Marlatte H (2017) Neurobiology of Schemas and Schema-Mediated Memory. Trends Cogn Sci 21:618–631.

Goddard E, Carlson TA, Dermody N, Woolgar A (2016) Representational dynamics of object recognition: Feedforward and feedback information flows. Neuroimage 128:385–397.

Goldberg II, Harel M, Malach R (2006) When the Brain Loses Its Self: Prefrontal Inactivation during Sensorimotor Processing. Neuron 50:329–339.

Golland Y, Golland P, Bentin S, Malach R (2008) Data-driven clustering reveals a fundamental subdivision of the human cortex into two global systems. Neuropsychologia 46:540–553.

Granger CWJ (1988) Causality, cointegration, and control. J Econ Dyn Control 12:551–559.

Gray WD (2007) Integrated models of cognitive systems. Oxford University Press.

Greene J, Haidt J (2002) How (and where) does moral judgment work? Trends Cogn Sci 6:517–523.

Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci 10:14–23.

Grill-Spector K, Malach R (2001) fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol (Amst) 107:293–321.

Grootswagers T, Wardle SG, Carlson TA (2016) Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time-series neuroimaging data. J Cogn Neurosci 29:677–697.

Grubert A, Eimer M (2018) The Time Course of Target Template Activation Processes during Preparation for Visual Search. J Neurosci 38:9527–9538.

Gusnard DA, Akbudak E, Shulman GL, Raichle ME (2001) Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. Proc Natl Acad Sci U S A 98:4259–4264.

Gusnard DA, Raichle ME (2001) Searching for a baseline: Functional imaging and the resting human brain. Nat Rev Neurosci 2:685–694.

Hahn B, Ross TJ, Stein EA (2007) Cingulate Activation Increases Dynamically with Response Speed under Stimulus Unpredictability. Cereb Cortex 17:1664–1671.

Han SW, Marois R (2013) Dissociation between process-based and data-based limitations for conscious perception in the human brain. Neuroimage 64:399–406.

Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. Nature 458:632–635.

Hartley T, Maguire EA, Spiers HJ, Burgess N (2003) The Well-Worn Route and the Path Less Traveled: Distinct Neural Bases of Route Following and Wayfinding in Humans. Neuron 37:877–888.

Hassabis D, Kumaran D, Maguire EA (2007) Using imagination to understand the neural basis of episodic memory. J Neurosci 27:14365–14374.

Hassabis D, Maguire EA (2007) Deconstructing episodic memory with construction. Trends Cogn Sci 11:299–306.

Hasson U, Chen J, Honey CJ (2015) Hierarchical process memory: Memory as an integral component of information processing. Trends Cogn Sci 19:304–313.

Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N (2008) A Hierarchy of Temporal Receptive Windows in Human Cortex. J Neurosci 28:2539–2550.

Hayama HR, Vilberg KL, Rugg MD (2012) Overlap between the Neural Correlates of Cued Recall and Source Memory: Evidence for a Generic Recollection Network? J Cogn Neurosci 24:1127–1137.

Hearne LJ, Mattingley JB, Cocchi L (2016) Functional brain networks related to individual differences in human intelligence at rest. Sci Rep 6:32328.

Hebart MN, Bankson BB, Harel A, Baker CI, Cichy RM (2018) The representational dynamics of task and object processing in humans. Elife 7.

Heller R, Golland Y, Malach R, Benjamini Y (2007) Conjunction group analysis: An alternative to mixed/random effect analysis. Neuroimage 37:1178–1185.

Honey CJ, Thesen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U (2012) Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. Neuron 76:423–434.

Hopf J-M, Boelmans K, Schoenfeld MA, Luck SJ, Heinze H-J (2004) Attention to features precedes attention to locations in visual search: evidence from electromagnetic brain responses in humans. J Neurosci 24:1822–1832.

Howard MW, Eichenbaum H (2013) The Hippocampus, Time, and Memory Across Scales.

Howard MW, Kahana MJ (2002) A Distributed Representation of Temporal Context. J Math Psychol 46:269–299.

Howard MW, Viskontas I V., Shankar KH, Fried I (2012) Ensembles of human MTL neurons "jump back in time" in response to a repeated stimulus. Hippocampus 22:1833–1847.

Howett D, Castegnaro A, Krzywicka K, Hagman J, Marchment D, Henson R, Rio M, King JA, Burgess N, Chan D (2019) Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation. Brain 142:1751–1766.

Hsieh LT, Gruber MJ, Jenkins LJ, Ranganath C (2014) Hippocampal Activity Patterns Carry Information about Objects in Temporal Context. Neuron 81:1165–1178.

Hsieh LT, Ranganath C (2015) Cortical and subcortical contributions to sequence retrieval: Schematic coding of temporal context in the neocortical recollection network. Neuroimage 121:78–90.

Humphreys GF, Hoffman P, Visser M, Binney RJ, Lambon Ralph MA (2015) Establishing task- and modality-dependent dissociations between the semantic and default mode networks. Proc Natl Acad Sci U S A 112:7857–7862.

Humphreys GF, Lambon Ralph MA (2017) Mapping domain-selective and counterpointed domain-general higher cognitive functions in the lateral parietal cortex: Evidence from fMRI comparisons of difficulty-varying semantic versus visuo-spatial tasks, and functional connectivity analyses. Cereb Cortex 27:4199–4212.

Isik L, Koldewyn K, Beeler D, Kanwisher N (2017) Perceiving social interactions in the posterior superior temporal sulcus. Proc Natl Acad Sci U S A 114:E9145–E9152.

Ito T, Kulkarni KR, Schultz DH, Mill RD, Chen RH, Solomyak LI, Cole MW (2017) Cognitive task information is transferred between brain regions via resting-state network topology. Nat Commun 8:1027.

Jackson J, Rich AN, Williams MA, Woolgar A (2017) Feature-selective Attention in Frontoparietal Cortex: Multivoxel Codes Adjust to Prioritize Task-relevant Information. J Cogn Neurosci 29:310–321.

James W (1890) The principles of psychology. Henry Holt and Company.

Jenkins LJ, Ranganath C (2010) Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. J Neurosci 30:15558–15565.

Jiang X, Bradley E, Rini RA, Zeffiro T, VanMeter J, Riesenhuber M (2007) Categorization Training Results in Shape- and Category-Selective Human Neural Plasticity. Neuron 53:891–903.

Johnson MK (2006) Memory and reality. Am Psychol 61:760–771.

Jovicich J, Peters RJ, Koch C, Braun J, Chang L, Ernst T (2001) Brain Areas Specific for Attentional Load in a Motion-Tracking Task. J Cogn Neurosci 13:1048–1058.

Kadohisa M, Petrov P, Stokes M, Sigala N, Buckley M, Gaffan D, Kusunoki M, Duncan J (2013) Dynamic Construction of a Coherent Attentional State in a Prefrontal Cell Population. Neuron 80:235–246.

Kalm K, Norris D (2014) The Representation of Order Information in Auditory-Verbal Short-Term Memory. J Neurosci 34:6879–6886.

Kalm K, Norris D (2017) Reading positional codes with fMRI: Problems and solutions de Lange FP, ed. PLoS One 12:e0176585.

Kan IP, Alexander MP, Verfaellie M (2009) Contribution of Prior Semantic Knowledge to New Episodic Learning in Amnesia. J Cogn Neurosci 21:938–944.

Kelley WM, Macrae CN, Wyland CL, Caglar S, Inati S, Heatherton TF (2002) Finding the Self? An Event-Related fMRI Study. J Cogn Neurosci 14:785–794.

Kim M, Maguire EA (2018) Hippocampus, Retrosplenial and Parahippocampal Cortices Encode Multicompartment 3D Space in a Hierarchical Manner. Cereb Cortex 28:1898–1909.

King JR, Dehaene S (2014) Characterizing the dynamics of mental representations: The temporal generalization method. Trends Cogn Sci 18:203–210.

Kiss M, Grubert A, Eimer M (2013) Top-down task sets for combined features: Behavioral and electrophysiological evidence for two stages in attentional object selection. Attention, Perception, Psychophys 75:216–228.

Koechlin E, Ody CC, Kouneiher FF (2003) The architecture of cognitive control in the human prefrontal cortex. Science 302:1181–1185.

Koechlin E, Summerfield C (2007) An information theoretical approach to prefrontal executive function. Trends Cogn Sci 11:229–235.

Kok P, Brouwer GJ, van Gerven MAJ, de Lange FP (2013) Prior Expectations Bias Sensory Representations in Visual Cortex. J Neurosci 33:16275–16284.

Kok P, Mostert P, de Lange FP (2017) Prior expectations induce prestimulus sensory templates. Proc Natl Acad Sci U S A 114:10473–10478.

Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping.

Proc Natl Acad Sci U S A 103:3863–3868.

Kucewicz MT, Berry BM, Kremen V, Brinkmann BH, Sperling MR, Jobst BC, Gross RE, Lega B, Sheth SA, Stein JM, Das SR, Gorniak R, Stead SM, Rizzuto DS, Kahana MJ, Worrell GA (2017) Dissecting gamma frequency activity during human memory processing. Brain 140:1337–1350.

Kuhl BA, Bainbridge WA, Chun MM (2012) Neural reactivation reveals mechanisms for updating memory. J Neurosci 32:3453–3461.

Kuhl BA, Rissman J, Chun MM, Wagner AD (2011) Fidelity of neural reactivation reveals competition between memories. Proc Natl Acad Sci U S A 108:5903–5908.

Kumaran D, Maguire EA (2006) An Unexpected Sequence of Events: Mismatch Detection in the Human Hippocampus Rugg M, ed. PLoS Biol 4:e424.

Kurby CA, Zacks JM (2008) Segmentation in the perception and memory of events. Trends Cogn Sci 12:72–79.

Laird AR, Fox PM, Eickhoff SB, Turner JA, Ray KL, McKay DR, Glahn DC, Beckmann CF, Smith SM, Fox PT (2011) Behavioral Interpretations of Intrinsic Connectivity Networks. J Cogn Neurosci 23:4022–4037.

Lashley KS (1951) The problem of serial order in behavior, Vol. 21. Bobbs-Merrill.

Lee S-H, Kravitz DJ, Baker CI (2013) Goal-dependent dissociation of visual and prefrontal cortices during working memory. Nat Neurosci 16:997–999.

Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011) Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. J Neurosci 31:2906–2915.

Levy BJ, Anderson MC (2002) Inhibitory processes and the control of memory retrieval. Trends Cogn Sci 6:299–305.

Lewis-Peacock JA, Norman KA (2014) Competition between items in working memory leads to forgetting. Nat Commun 5:5768.

Li S, Ostwald D, Giese M, Kourtzi Z (2007) Flexible Coding for Categorical Decisions in the Human Brain. J Neurosci 27:12321–12330.

Liesefeld HR (2018) Estimating the Timing of Cognitive Operations With MEG/EEG Latency Measures: A Primer, a Brief Tutorial, and an Implementation of Various Methods. Front Neurosci 12:765.

Linden DEJ, Bittner RA, Muckli L, Waltz JA, Kriegeskorte N, Goebel R, Singer W, Munk MHJ (2003) Cortical capacity constraints for visual working memory: Dissociation of fMRI load effects in a fronto-parietal network. Neuroimage 20:1518–1530.

Logie RH, Law A, Trawley S, Nissan J (2010) Multitasking, working memory and

remembering intentions. Psychol Belg 50:309.

Logie RH, Trawley S, Law A (2011) Multitasking: multiple, domain-specific cognitive functions in a virtual environment. Mem Cognit 39:1561–1574.

Luck SJ (Steven J, Kappenman ES, Oxford University Press. (2012) Oxford handbook of event-related potential components. Oxford University Press.

Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997) Neural Mechanisms of Spatial Selective Attention in Areas V1, V2, and V4 of Macaque Visual Cortex. J Neurophysiol 77:24–42.

Luck SJ, Hillyard SA (1994) Spatial filtering during visual search: evidence from human electrophysiology. J Exp Psychol Hum Percept Perform 20:1000–1014.

Lundqvist M, Compte A, Lansner A (2010) Bistable, Irregular Firing and Population Oscillations in a Modular Attractor Memory Network Morrison A, ed. PLoS Comput Biol 6:e1000803.

Lundqvist M, Herman P, Miller EK (2018) Working Memory: Delay Activity, Yes! Persistent Activity? Maybe Not. J Neurosci 38:7013–7019.

Lundqvist M, Rose J, Herman P, Brincat SL, Buschman TJ, Miller EK (2016a) Gamma and Beta Bursts Underlie Working Memory. Neuron 90:152–164.

Lundqvist M, Rose J, Herman P, Brincat SL, Buschman TJ, Miller EK (2016b) Gamma and Beta Bursts Underlie Working Memory. Neuron 90:152–164.

Luria R, Balaban H, Awh E, Vogel EK (2016) The contralateral delay activity as a neural measure of visual working memory. Neurosci Biobehav Rev 62.

Maguire EA (2012) Studying the freely-behaving brain with fMRI. Neuroimage 62:1170–1176.

Maguire EA, Burgess N, Donnett JG, Frackowiak RS, Frith CD, O'Keefe J (1998) Knowing where and getting there: a human navigation network. Science 280:921–924.

Manly T, Hawkins K, Evans J, Woldt K, Robertson IH (2002) Rehabilitation of executive function: Facilitation of effective goal management on complex tasks using periodic auditory alerts. Neuropsychologia 40:271–281.

Manning JR (2019) Episodic memory: mental time travel or a quantum `memory wave' function? PsyArXiv.

Manning JR, Hulbert JC, Williams J, Piloto L, Sahakyan L, Norman KA (2016) A neural signature of contextually mediated intentional forgetting. Psychon Bull Rev 23:1534–1542.

Manning JR, Kahana MJ, Norman KA (2014) The role of context in episodic memory. In: The Cognitive Neurosciences, 5th ed. (Gazzaniga M E, ed). Cambridge, MA: MIT Press.

Manning JR, Polyn SM, Baltuch GH, Litt B, Kahana MJ (2011) Oscillatory patterns in temporal

lobe reveal context reinstatement during memory search. Proc Natl Acad Sci U S A 108:12893–12897.

Manoach DS, Schlaug G, Siewert B, Darby DG, Bly BM, Benfield A, Edelman RR, Warach S (1997) Prefrontal cortex fMRI signal changes are correlated with working memory load. Neuroreport 8:545–549.

Mantini D, Vanduffel W (2013) Emerging Roles of the Brain's Default Network. Neurosci 19:76–87.

Marois R, Chun MM, Gore JC (2004) A Common Parieto-Frontal Network Is Recruited Under Both Low Visibility and High Perceptual Interference Conditions. J Neurophysiol 92:2985–2992.

Marois R, Ivanoff J (2005) Capacity limits of information processing in the brain. Trends Cogn Sci 9:296–305.

Mars RB, Neubert F-X, Noonan MP, Sallet J, Toni I, Rushworth MFS (2012) On the relationship between the "default mode network" and the "social brain." Front Hum Neurosci 6:189.

Martuzzi R, Ramani R, Qiu M, Rajeevan N, Constable RT (2010) Functional connectivity and alterations in baseline brain state in humans. Neuroimage 49:823–834.

Mason MF, Norton MI, Horn JD Van, Wegner DM, Grafton ST, Macrae CN (2007) Wandering Minds: The Default Network and Stimulus-Independent Thought. Science (80-) 315:393–395.

Mazoyer B, Zago L, Mellet E, Bricogne S, Etard O, Houdé O, Crivello F, Joliot M, Petit L, Tzourio-Mazoyer N (2001) Cortical networks for working memory and executive functions sustain the conscious resting state in man.

Mazoyer P, Wicker B, Fonlupt P (2002) A neural network elicited by parametric manipulation of the attention load. Neuroreport 13:2331–2334.

McDermott KB, Szpunar KK, Christ SE (2009) Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. Neuropsychologia 47:2290–2298.

McGeorge P, Phillips LH, Crawford JR, Garden SE, Sala S Della, Milne AB, Hamilton S, Callender JS (2001) Using Virtual Environments in the Assessment of Executive Dysfunction. Presence Teleoperators Virtual Environ 10:375–383.

McGuire PK, Paulesu E, Frackowiak RS, Frith CD (1996) Brain activity during stimulus independent thought. Neuroreport 7:2095–2099.

McKiernan KA, D'Angelo BR, Kaufman JN, Binder JR (2006) Interrupting the "stream of

consciousness": An fMRI investigation. Neuroimage 29:1185–1191.

McKiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR (2003) A Parametric Manipulation of Factors Affecting Task-induced Deactivation in Functional Neuroimaging. J Cogn Neurosci 15:394–408.

McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser M-B (2006) Path integration and the neural basis of the "cognitive map." Nat Rev Neurosci 7:663–678.

Menon V (2011) Large-scale brain networks and psychopathology: a unifying triple network model. Trends Cogn Sci 15:483–506.

Meyer T, Qi X-L, Stanford TR, Constantinidis C (2011) Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. J Neurosci 31:6266–6276.

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. :167–202.

Miller EK, Lundqvist M, Bastos AM (2018) Working Memory 2.0. Neuron 100:463–475.

Mitchell DJ, Cusack R (2008) Flexible, Capacity-Limited Activity of Posterior Parietal Cortex in Perceptual as well as Visual Short-Term Memory Tasks. Cereb Cortex 18:1788–1798.

Mitchell KJ, Johnson MK (2009) Source Monitoring 15 Years Later: What Have We Learned From fMRI About the Neural Mechanisms of Source Memory?

Molenberghs P, Johnson H, Henry JD, Mattingley JB (2016) Understanding the minds of others: A neuroimaging meta-analysis. Neurosci Biobehav Rev 65:276–291.

Moll J, Zahn R, de Oliveira-Souza R, Krueger F, Grafman J (2005) The neural basis of human moral cognition. Nat Rev Neurosci 6:799–809.

Moore M, Hu Y, Woo S, O'Hearn D, Iordan AD, Dolcos S, Dolcos F (2014) A Comprehensive Protocol for Manual Segmentation of the Medial Temporal Lobe Structures. J Vis Exp:1–8.

Muhle-Karbe PS, Duncan J, De Baene W, Mitchell DJ, Brass M (2016) Neural Coding for Instruction-Based Task Sets in Human Frontoparietal and Visual Cortex. Cereb Cortex 27:bhw032.

Muller-Gass A, Schröger E (2007) Perceptual and cognitive task difficulty has differential effects on auditory distraction. Brain Res 1136:169–177.

Murphy C, Jefferies E, Rueschemeyer S-A, Sormaz M, Wang H, Margulies DS, Smallwood J (2018) Distant from input: Evidence of regions within the default mode network supporting perceptually-decoupled and conceptually-guided cognition. Neuroimage 171:393–401.

Myers NE, Rohenkohl G, Wyart V, Woolrich MW, Nobre AC, Stokes MG (2015) Testing

sensory evidence against mnemonic templates. Elife 4:1–25.

Niendam TA, Laird AR, Ray KL, Dean YM, Glahn DC, Carter CS (2012) Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. Cogn Affect Behav Neurosci 12:241–268.

Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N (2014) A Toolbox for Representational Similarity Analysis Prlic A, ed. PLoS Comput Biol 10:e1003553.

Noonan KA, Jefferies E, Visser M, Lambon Ralph MA (2013) Going beyond Inferior Prefrontal Involvement in Semantic Control: Evidence for the Additional Contribution of Dorsal Angular Gyrus and Posterior Middle Temporal Cortex. J Cogn Neurosci 25:1824–1850.

Norman DA, Bobrow DG (1975) On data-limited and resource-limited processes. Cogn Psychol 7:44–64.

O'Craven KM, Downing PE, Kanwisher N (1999) fMRI evidence for objects as the units of attentional selection. Nat 1999 4016753 401:584.

O'Reilly RC, Rudy JW (2001) Conjunctive representations in learning and memory: principles of cortical and hippocampal function. Psychol Rev 108:311–345.

Olivers CNL, Peters J, Houtkamp R, Roelfsema PR (2011) Different states in visual working memory: When it guides attention and when it does not. Trends Cogn Sci 15:327–334.

Olsson A, Ochsner KN (2008) The role of social cognition in emotion. Trends Cogn Sci 12:65–71.

Parthasarathy A, Herikstad R, Bong JH, Medina FS, Libedinsky C, Yen S-C (2017) Mixed selectivity morphs population codes in prefrontal cortex. Nat Neurosci 20:1770–1779.

Peelen M V., Fei-Fei L, Kastner S (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. Nature 460:94–97.

Peelen M V., Kastner S (2011) A neural basis for real-world visual search in human occipitotemporal cortex. Proc Natl Acad Sci 108:12125–12130.

Penfield W, Evans J (1935) The frontal lobe in man: a clinical study of maximum removals. Brain 58:115–133.

Polyn SM, Kahana MJ (2008) Memory search and the neural representation of context. Trends Cogn Sci 12:24–30.

Polyn SM, Natu VS, Cohen JD, Norman KA (2005) Category-specific cortical activity precedes retrieval during memory search. Science 310:1963–1966.

Polyn SM, Norman KA, Kahana MJ (2008) A context maintenance and retrieval model of organizational processes in free recall.

Polyn SM, Norman KA, Kahana MJ (2009) Task context and organization in free recall.

Neuropsychologia 47:2158–2163.

Postle BR (2006) Working memory as an emergent property of the mind and brain. Neuroscience 139:23–38.

Preston AR, Eichenbaum H (2013) Interplay of Hippocampus and Prefrontal Cortex in Memory. Curr Biol 23:R764–R773.

Puri AM, Wojciulik E, Ranganath C (2009) Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. Brain Res 1301:89–99.

Radvansky GA, Zacks JM (2011) Event perception. Wiley Interdiscip Rev Cogn Sci 2:608–620.

Radvansky GA, Zacks JM (2017) Event boundaries in memory and cognition. Curr Opin Behav Sci 17:133–140.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. Proc Natl Acad Sci 98:676–682.

Ranganath C (2010a) A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. Hippocampus.

Ranganath C (2010b) Binding Items and Contexts: The Cognitive Neuroscience of Episodic Memory Medial Temporal Lobes and Binding of Items and Contexts. Curr Dir Psychol Sci.

Ranganath C, Ritchey M (2012) Two cortical systems for memory-guided behaviour. Nat Rev Neurosci.

Reagh ZM, Ranganath C (2018) What does the functional organization of cortico-hippocampal networks tell us about the functional organization of memory? Neurosci Lett 680:69–76.

Reverberi C, Gorgen K, Haynes J-D (2012a) Distributed Representations of Rule Identity and Rule Order in Human Frontal Cortex and Striatum. J Neurosci 32:17420–17430.

Reverberi C, Görgen K, Haynes J-D (2012b) Compositionality of Rule Representations in Human Prefrontal Cortex. Cereb Cortex 22:1237–1246.

Reynolds JR, Zacks JM, Braver TS (2007) A Computational Model of Event Segmentation From Perceptual Prediction. Cogn Sci 31:613–643.

Richmond LL, Zacks JM (2017) Constructing Experience: Event Models from Perception to Action. Trends Cogn Sci 21:962–980.

Rilling JK, Sanfey AG, Aronson JA, Nystrom LE, Cohen JD (2004) The neural correlates of theory of mind within interpersonal interactions. Neuroimage 22:1694–1703.

Ritchey M, Montchal ME, Yonelinas AP, Ranganath C (2015) Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. Elife 4.

Robin J, Moscovitch M (2017) Details, gist and schema: hippocampal–neocortical interactions underlying recent and remote episodic and spatial memory. Curr Opin Behav Sci 17:114–123.

Roca M, Manes F, Chade A, Gleichgerrcht E, Gershanik O, Arévalo GG, Torralva T, Duncan J (2012) The relationship between executive functions and fluid intelligence in Parkinson's disease. Psychol Med 42:2445–2452.

Roca M, Parr A, Thompson R, Woolgar A, Torralva T, Antoun N, Manes F, Duncan J (2009) Executive function and fluid intelligence after frontal lobe lesions. Brain 133:234–247.

Roca M, Torralva T, Gleichgerrcht E, Woolgar A, Thompson R, Duncan J, Manes F (2011) The role of Area 10 (BA10) in human multitasking and in social cognition: A lesion study. Neuropsychologia 49:3525–3531.

Rogers RD, Monsell S (1995) Costs of a predictible switch between simple cognitive tasks. J Exp Psychol Gen 124:207–231.

Rosenbaum DA, Kenny SB, Derr MA (1983) Hierarchical control of rapid movement sequences. J Exp Psychol Hum Percept Perform 9:86–102.

Ross RS, Brown TI, Stern CE (2009) The retrieval of learned sequences engages the hippocampus: Evidence from fMRI. Hippocampus 19:790–799.

Rothbart MK, Posner MI (2015) The developing brain in a multitasking world. Dev Rev 35:42–63.

Rugg MD, Vilberg KL (2013) Brain networks underlying episodic memory retrieval. Curr Opin Neurobiol 23:255–260.

Rugg MD, Vilberg KL, Mattson JT, Yu SS, Johnson JD, Suzuki M (2012) Item memory, context memory and the hippocampus: fMRI evidence. Neuropsychologia 50:3070–3079.

Saenz M, Buracas GT, Boynton GM (2002) Global effects of feature-based attention in human visual cortex. Nat Neurosci 5:631–632.

Saga Y, Iba M, Tanji J, Hoshi E (2011) Development of multidimensional representations of task phases in the lateral prefrontal cortex. J Neurosci 31:10648–10665.

Sämann PG, Wehrle R, Hoehn D, Spoormaker VI, Peters H, Tully C, Holsboer F, Czisch M (2011) Development of the Brain's Default Mode Network from Wakefulness to Slow Wave Sleep. Cereb Cortex 21:2082–2093.

Saxe R, Kanwisher N (2003) People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind." Neuroimage 19:1835–1842.

Schacter DL, Addis DR (2007) The cognitive neuroscience of constructive memory: remembering the past and imagining the future. Philos Trans R Soc B Biol Sci 362:773–

786.

Schacter DL, Addis DR, Buckner RL (2008) Episodic Simulation of Future Events. Ann N Y Acad Sci 1124:39–60.

Schiller D, Eichenbaum H, Buffalo EA, Davachi L, Foster DJ, Leutgeb S, Ranganath C (2015) Memory and Space: Towards an Understanding of the Cognitive Map. J Neurosci 35:13904–13911.

Schneider DW, Logan GD (2006) Hierarchical control of cognitive processes: Switching tasks in sequences. J Exp Psychol Gen 135:623–640.

Schoenfeld MA, Tempelmann C, Martinez A, Hopf J-M, Sattler C, Heinze H-J, Hillyard SA (2003) Dynamics of feature binding during object-selective attention. Proc Natl Acad Sci U S A 100:11806–11811.

Seidl KN, Peelen M V., Kastner S (2012) Neural Evidence for Distracter Suppression during Visual Search in Real-World Scenes. J Neurosci 32:11812–11819.

Serences J, Ester E, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. Psychol Sci 20:207–214.

Serences JT, Boynton GM (2007) Feature-Based Attentional Modulations in the Absence of Direct Visual Stimulation. Neuron 55:301–312.

Sergent C, Baillet S, Dehaene S (2005) Timing of the brain events underlying access to consciousness during the attentional blink. Nat Neurosci 8:1391–1400.

Seth AK, Barrett AB, Barnett L (2015) Granger Causality Analysis in Neuroscience and Neuroimaging. J Neurosci 35:3293–3297.

Shafi M, Zhou Y, Quintana J, Chow C, Fuster J, Bodner M (2007) Variability in neuronal activity in primate cortex during working memory tasks. Neuroscience 146:1082–1108.

Shallice T, Burgess PW (1991) Deficits in Strategy Application. Brain:727–741.

Shibata K, Yamagishi N, Goda N, Yoshioka T, Yamashita O, Sato M, Kawato M (2008) The Effects of Feature Attention on Prestimulus Cortical Activity in the Human Visual System. Cereb Cortex 18:1664–1675.

Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE (1997) Common Blood Flow Changes across Visual Tasks: II. Decreases in Cerebral Cortex. J Cogn Neurosci 9:648–663.

Sigala N, Kusunoki M, Nimmo-Smith I, Gaffan D, Duncan J (2008) Hierarchical coding for sequential task events in the monkey prefrontal cortex. Proc Natl Acad Sci U S A 105:11969–11974.

Silva M, Baldassano C, Fuentemilla L (2019) Rapid memory reactivation at movie event

boundaries promotes episodic encoding. bioRxiv:511782.

Simpson G V., Weber DL, Dale CL, Pantazis D, Bressler SL, Leahy RM, Luks TL (2011) Dynamic Activation of Frontal, Parietal, and Sensory Regions Underlying Anticipatory Visual Spatial Attention. J Neurosci 31:13880–13889.

Small D., Gitelman D., Gregory M., Nobre A., Parrish T., Mesulam M-M (2003) The posterior cingulate and medial prefrontal cortex mediate the anticipatory allocation of spatial attention. Neuroimage 18:633–641.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF (2009) Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci 106:13040–13045.

Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44:83–98.

Smith SM, Vela E (2001) Environmental context-dependent memory: A review and meta-analysis. Psychon Bull Rev 8:203–220.

Soon CS, Namburi P, Chee MWL (2013) Preparatory patterns of neural activity predict visual category search speed. Neuroimage 66:215–222.

Spalding KN, Jones SH, Duff MC, Tranel D, Warren DE (2015) Investigating the Neural Correlates of Schemas: Ventromedial Prefrontal Cortex Is Necessary for Normal Schematic Influence on Memory. J Neurosci 35:15746–15751.

Speer NK, Zacks JM, Reynolds JR (2007) Human brain activity time-locked to narrative event boundaries. Psychol Sci 18:449–455.

Spiers HJ, Gilbert SJ (2015) Solving the detour problem in navigation: a model of prefrontal and hippocampal interactions. Front Hum Neurosci 9:125.

Spiers HJ, Maguire EA (2006) Thoughts, behaviour, and brain dynamics during navigation in the real world. Neuroimage 31:1826–1840.

Spiers HJ, Maguire EA (2007) Decoding human brain activity during real-world experiences. Trends Cogn Sci 11:356–365.

Spreng RN (2012) The Fallacy of a "Task-Negative" Network. Front Psychol 3:145.

Spreng RN, Grady CL (2010) Patterns of Brain Activity Supporting Autobiographical Memory, Prospection, and Theory of Mind, and Their Relationship to the Default Mode Network. J Cogn Neurosci 22:1112–1123.

Spreng RN, Mar RA, Kim ASN (2009) The Common Neural Basis of Autobiographical Memory, Prospection, Navigation, Theory of Mind, and the Default Mode: A Quantitative

Meta-analysis. J Cogn Neurosci 21:489–510.

Spreng RN, Stevens WD, Chamberlain JP, Gilmore AW, Schacter DL (2010) Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. Neuroimage 53:303–317.

Sridharan D, Levitin DJ, Chafe CH, Berger J, Menon V (2007) Neural Dynamics of Event Segmentation in Music: Converging Evidence for Dissociable Ventral and Dorsal Networks. Neuron 55:521–532.

Stawarczyk D, Majerus S, Maquet P, D'Argembeau A (2011) Neural Correlates of Ongoing Conscious Experience: Both Task-Unrelatedness and Stimulus-Independence Are Related to Default Network Activity Gilbert S, ed. PLoS One 6:e16997.

Stokes M, Spaak E (2016) The Importance of Single-Trial Analyses in Cognitive Neuroscience. Trends Cogn Sci 20:483–486.

Stokes M, Thompson R, Nobre AC, Duncan J (2009) Shape-specific preparatory activity mediates attention to targets in human visual cortex. Proc Natl Acad Sci 106:19569–19574.

Stokes MG (2015) "Activity-silent" working memory in prefrontal cortex: A dynamic coding framework. Trends Cogn Sci 19:394–405.

Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J (2013) Dynamic coding for cognitive control in prefrontal cortex. Neuron 78:364–375.

Stokes MG, Wolff MJ, Spaak E (2015) Decoding Rich Spatial Information with High Temporal Resolution. Trends Cogn Sci 19:636–638.

Svoboda E, McKinnon MC, Levine B (2006) The functional neuroanatomy of autobiographical memory: A meta-analysis. Neuropsychologia 44:2189–2208.

Swallow KM, Barch DM, Head D, Maley CJ, Holder D, Zacks JM (2011) Changes in Events Alter How People Remember Recent Information. J Cogn Neurosci 23:1052–1064.

Sylvester CM, Shulman GL, Jack AI, Corbetta M (2009) Anticipatory and Stimulus-Evoked Blood Oxygenation Level-Dependent Modulations Related to Spatial Attention Reflect a Common Additive Signal. J Neurosci 29:10671–10682.

Tavares P, Lawrence AD, Barnard PJ (2008) Paying Attention to Social Meaning: An fMRI Study. Cereb Cortex 18:1876–1885.

Tavares RM, Mendelsohn A, Grossman Y, Williams CH, Shapiro M, Trope Y, Schiller D (2015) A Map for Social Navigation in the Human Brain. Neuron 87:231–243.

Todd JJ, Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. Nature 428:751–754.

Tolman EC (1948) Cognitive maps in rats and men. Psychol Rev 55:189–208.

Treue S, Martínez Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399:575–579.

Tschentscher N, Mitchell D, Duncan J (2017) Fluid Intelligence Predicts Novel Rule Implementation in a Distributed Frontoparietal Control Network. J Neurosci 37:4841–4847.

Tulving E (1972) Episodic and semantic memory. In: Organization of memory (Tulving, E; Donaldson W, ed). Oxford, England: Academic Press.

Tulving E (2002) Episodic Memory: From Mind to Brain. Annu Rev Psychol 53:1–25.

van Driel J, Olivers CN., Fahrenfort JJ (2019) High-pass filtering artifacts in multivariate classification of neural time series data. bioRxiv:530220.

van Kesteren MTR, Beul SF, Takashima A, Henson RN, Ruiter DJ, Fernández G (2013) Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. Neuropsychologia 51:2352–2359.

van Kesteren MTR, Ruiter DJ, Fernández G, Henson RN (2012) How schema and novelty augment memory formation. Trends Neurosci 35:211–219.

Van Overwalle F (2011) A dissociation between social mentalizing and general reasoning. Neuroimage 54:1589–1599.

Vann SD, Aggleton JP, Maguire EA (2009) What does the retrosplenial cortex do? Nat Rev Neurosci 10:792–802.

Vidaurre D, Myers NE, Stokes M, Nobre AC, Woolrich MW (2019) Temporally Unconstrained Decoding Reveals Consistent but Time-Varying Stages of Stimulus Processing. Cereb Cortex 29:863–874.

Vogel EK, Machizawa MG (2004) Neural activity predicts individual differences in visual working memory capacity. Nature 428:748–751.

Vogt BA, Finch DM, Olson CR (1992) Functional Heterogeneity in Cingulate Cortex: The Anterior Executive and Posterior Evaluative Regions. Cereb Cortex 2:435–443.

Wager TD, Smith EE (2003) Neuroimaging studies of working memory: a meta-analysis. Cogn Affect & Behav Neurosci 3:255–274.

Wallis G, Stokes M, Cousijn H, Woolrich M, Nobre AC (2015) Frontoparietal and Cingulo-opercular Networks Play Dissociable Roles in Control of Working Memory. J Cogn Neurosci 27:2019–2034.

Walther A, Nili H, Ejaz N, Alink A, Kriegeskorte N, Diedrichsen J (2016) Reliability of dissimilarity measures for multi-voxel pattern analysis. Neuroimage 137:188–200.

Wang W-C, Yonelinas AP, Ranganath C (2013) Dissociable neural correlates of item and context retrieval in the medial temporal lobes. Behav Brain Res 254:102–107.

Wimber M, Alink A, Charest I, Kriegeskorte N, Anderson MC (2015) Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. Nat Neurosci 18:582–589.

Wolfe JM (1994) Guided Search 2.0 A revised model of visual search. Psychon Bull Rev 1:202–238.

Wolff MJ, Ding J, Myers NE, Stokes MG (2015) Revealing hidden states in visual working memory using electroencephalography. Front Syst Neurosci 9:1–12.

Wolff MJ, Jochim J, Akyürek EG, Stokes MG (2017) Dynamic hidden states underlying working-memory-guided behavior. Nat Neurosci.

Woolgar A, Hampshire A, Thompson R, Duncan J (2011a) Adaptive Coding of Task-Relevant Information in Human Frontoparietal Cortex. J Neurosci 31:14592–14599.

Woolgar A, Thompson R, Bor D, Duncan J (2011b) Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. Neuroimage 56:744–752.

Woolgar A, Williams MA, Rich AN (2015) Attention enhances multi-voxel representation of novel objects in frontal, parietal and visual cortices. Neuroimage 109:429–437.

Wutz A, Loonis R, Roy JE, Donoghue JA, Miller EK (2018) Different Levels of Category Abstraction by Different Dynamics in Different Prefrontal Areas. Neuron.

Wylie G, Allport A (2000) Task switching and the measurement of "switch costs" Psychol Res 63:212–233.

Xu Y, Chun MM (2009) Selecting and perceiving multiple visual objects. Trends Cogn Sci 13:167–174.

Yarkoni T, Speer NK, Zacks JM (2008) Neural substrates of narrative comprehension and memory. Neuroimage 41:1408–1425.

Yeo BTT, Krienen FM, Eickhoff SB, Yaakub SN, Fox PT, Buckner RL, Asplund CL, Chee MWL (2015) Functional specialization and flexibility in human association cortex. Cereb Cortex 25:3654–3672.

Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125–1165.

Yeshurun Y, Nguyen M, Hasson U (2017) Amplification of local changes along the timescale processing hierarchy. Proc Natl Acad Sci U S A 114:9475–9480.

Young L, Dodell-Feder D, Saxe R (2010) What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. Neuropsychologia 48:2658–2664.

Zacks JM, Braver TS, Sheridan M a, Donaldson DI, Snyder AZ, Ollinger JM, Buckner RL, Raichle ME (2001a) Human brain activity time-locked. :651–655.

Zacks JM, Speer NK, Swallow KM, Braver TS, Reynolds JR (2007) Event perception: A mind-brain perspective. Psychol Bull 133:273–293.

Zacks JM, Speer NK, Vettel JM, Jacoby LL (2006) Event Understanding and Memory in Healthy Aging and Dementia of the Alzheimer Type.

Zacks JM, Swallow KM (2007) Event Segmentation. Curr Dir Psychol Sci 16:80–84.

Zacks JM, Tversky B (2001) Event structure in perception and conception. Psychol Bull 127:3–21.

Zacks JM, Tversky B, Iyer G (2001b) Perceiving, remembering, and communicating structure in events. J Exp Psychol Gen 130:29–58.

Zadbood A, Chen J, Leong YC, Norman KA, Hasson U (2017) How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. Cereb Cortex 27:4988–5000.

Zhang W, Luck SJ (2009) Feature-based attention modulates feedforward visual processing. Nat Neurosci 12:24–25.

Zwaan RA, Radvansky GA (1998) Situation models in language comprehension and memory. Psychol Bull 123:162–185.

Appendices

Appendix A. Supplementary materials for Chapter 2

Since complex objects were used in the experiment, and higher visual regions are specialized for object-level processing, we additionally examined a broad extrastriate visual cortex (ESV) ROI from the Fedorenko et al. (2013) template, which encompasses object, face, and scene processing regions. In all cases, results were very similar to those from V1, presumably reflecting the relatively low resolution of EEG/MEG. We report the results from ESV here.

Coding of the attentional cue/attentional template during the preparatory phase

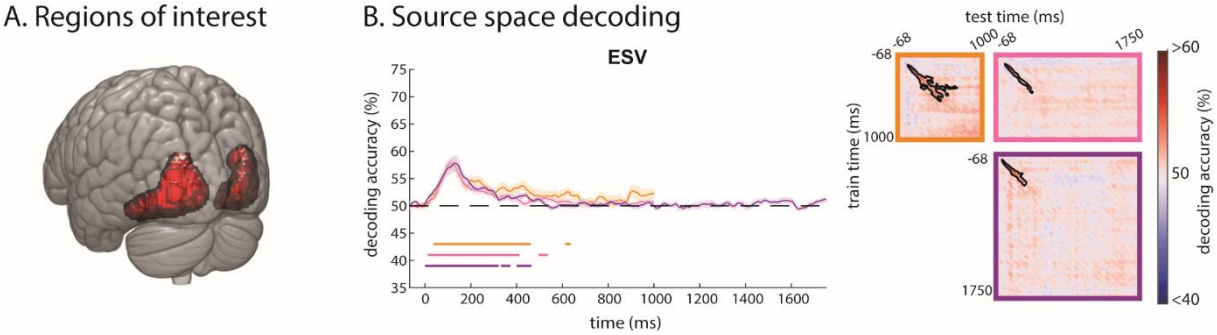


Figure A.1. (A) Vertices within source space ESV ROI. (B) Decoding time-course of auditory stimulus/attentional cue from ESV ROI. Curves on the left show decoding when training and testing on matched time-points. Colored dots beneath the decoding curves show times where decoding on the diagonal is significantly above chance for each condition ($p < 0.05$). Translucent bands represent standard error of the mean. Matrices on the right show temporal generalization of decoding across all pairs of training and testing times. Black contours indicate regions of significant decoding ($p<0.05$). Significance is corrected for multiple comparisons across *time using TFCE and permutation testing.*

To test whether activity during any stage of the preparatory phase might reflect the representation of the upcoming trial target, we performed a cross-task and cross-time classification analysis trained using the visual localizer task. Replicating the results of the three ROIs used in the main experiment, we did not find any significant clusters where the visual template cross-generalized to the preparatory phase in ESV.

Coding of visual properties of 1-item displays

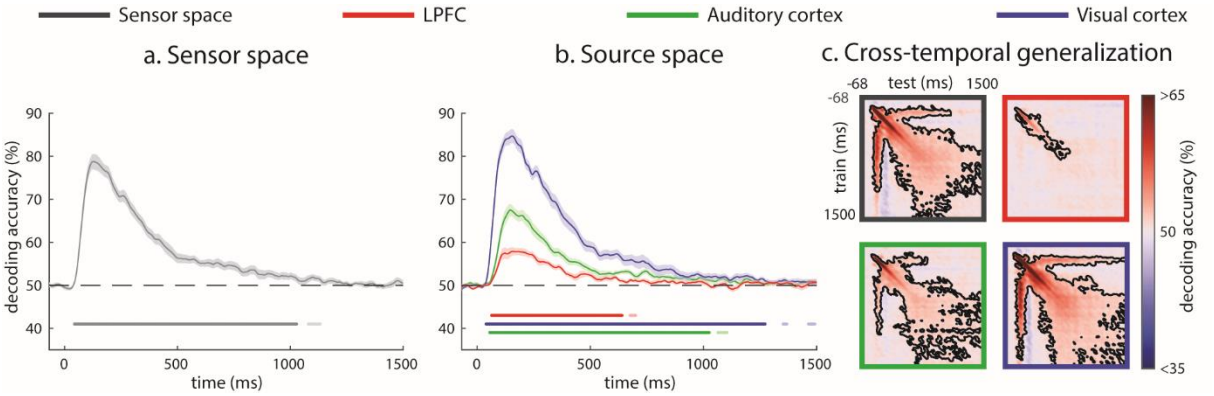


Figure A.2. Coding of visual properties of 1-item displays. Decoding time-courses of object identity, in (a) sensor and (b) source space, when training/testing using matched time-points, and (c) generalizing across training/testing times. Dark colored dots beneath the decoding curves show times where decoding is significantly above chance for each condition ($p < 0.05$), corrected for multiple comparisons along the diagonal of the cross-temporal generalization matrix; faint colored dots represent additional time-points where the diagonal of the cross-temporal generalization matrix is significant when corrected for multiple comparisons across the whole matrix. Translucent bands represent standard error of the mean. Matrices on the right show temporal generalization of decoding across all pairs of training and testing times. Black contours indicate regions of significant decoding ($p<0.05$). Significance is corrected for multiple comparisons across time using TFCE and permutation testing.

Coding of behavioral properties of 1-item displays

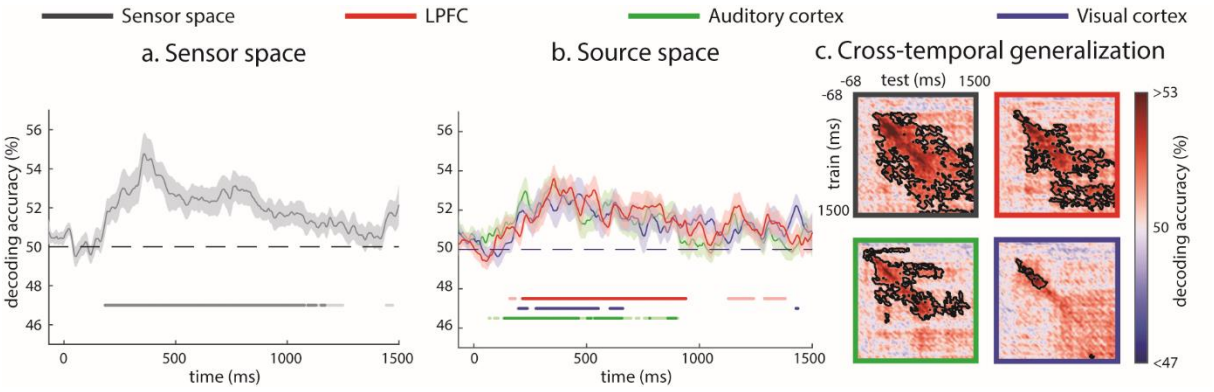


Figure A.3. Coding of behavioral properties of 1-item displays. Same format as figure above.

Coding of target location in 3-item displays

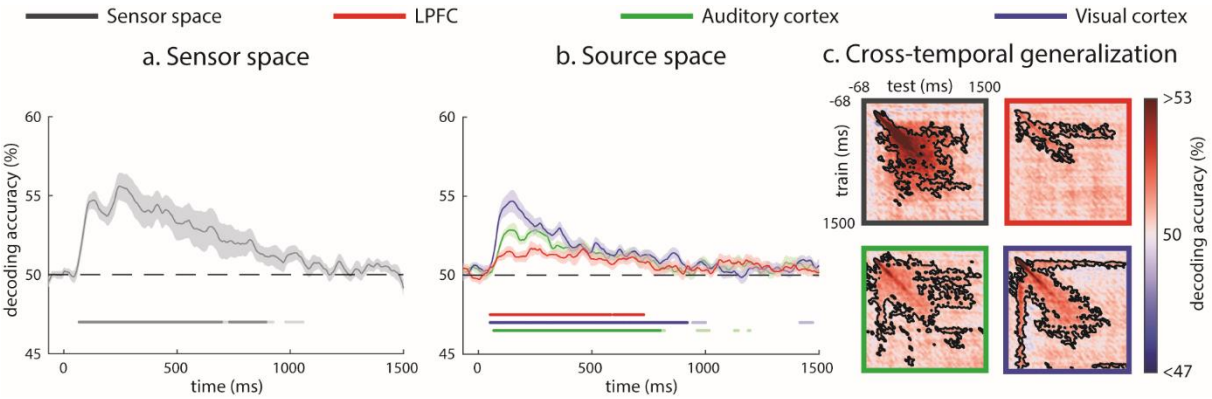


Figure A.4. Coding of target location in 3-item displays. Same format as figure above.

Coding of target identity during presentation of 3-item displays

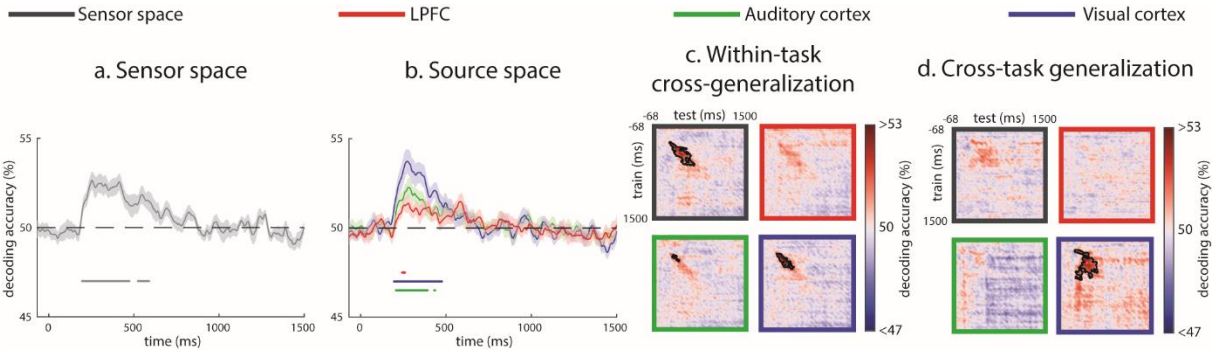


Figure A.5. Decoding of attentional cue/target identity during presentation of 3-item displays. Same format as figure above.

Reawakening of the attentional cue/template during presentation of consistent non-targets

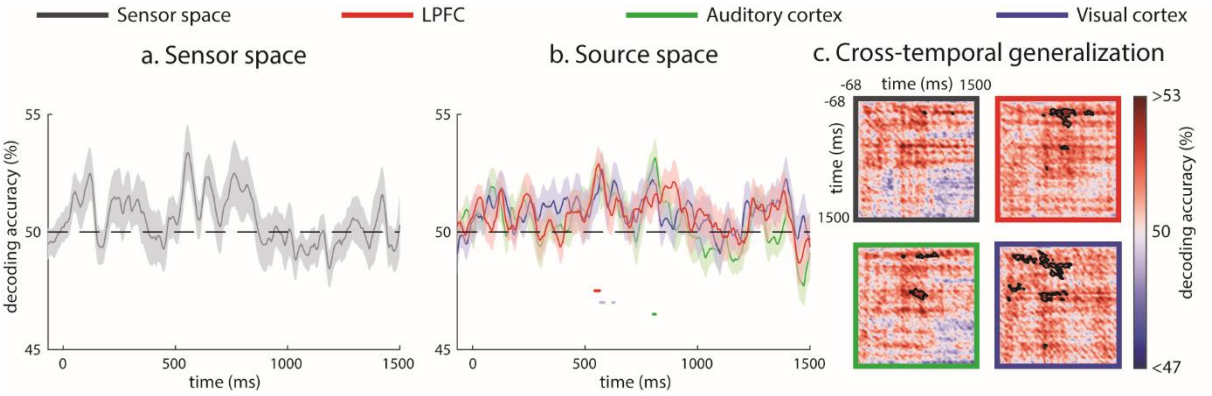


Figure A.6. Decoding time-course of attention cue during presentation of Nc displays. Same format as figure above.

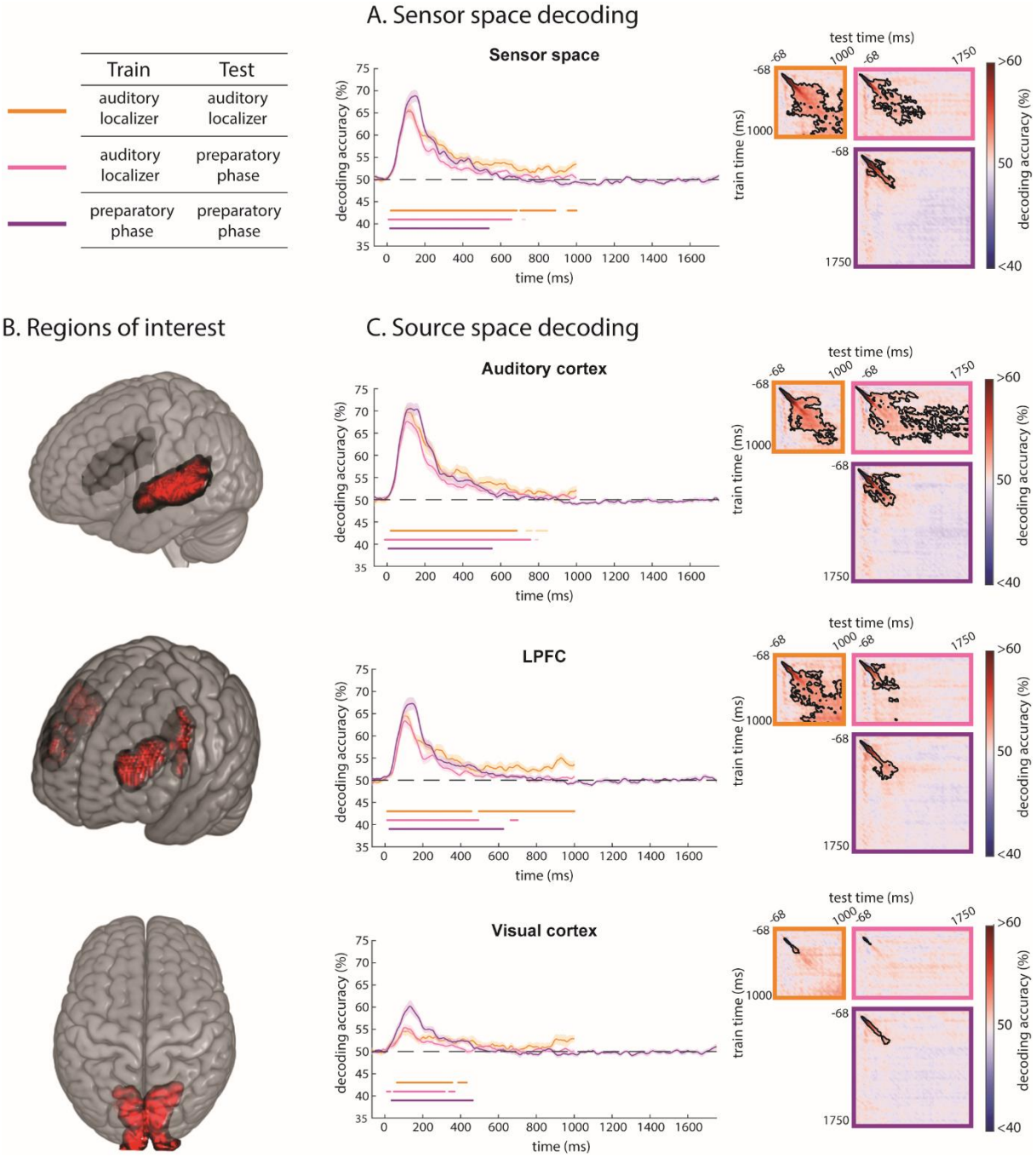


Figure A.7. Response to the attentional cue without subsampling the preparatory phase data. (A) Decoding time-course of auditory stimulus/attentional cue using all sensors combining EEG and MEG across the whole brain. Curves on the left show decoding when training and testing on matched time-points. Dark colored dots beneath the decoding curves show times where decoding is significantly above chance for each condition ($p < 0.05$), corrected for multiple comparisons along the diagonal of the cross-temporal generalization matrix; faint colored dots represent additional time-points where the diagonal of the cross-temporal generalization matrix is significant when corrected for multiple comparisons across the whole matrix Translucent bands represent standard error of the mean. Matrices on the right show

temporal generalization of decoding across all pairs of training and testing times. Black contours indicate regions of significant decoding ($p<0.05$). (B) Vertices within source space ROIs (auditory cortex, lateral prefrontal cortex (LPFC), and visual cortex). (C) Decoding time-courses from these source space ROIs; same format as (A). Significance is corrected for multiple comparisons across time using TFCE and permutation testing.

Appendix B. Supplementary materials for Chapter 3

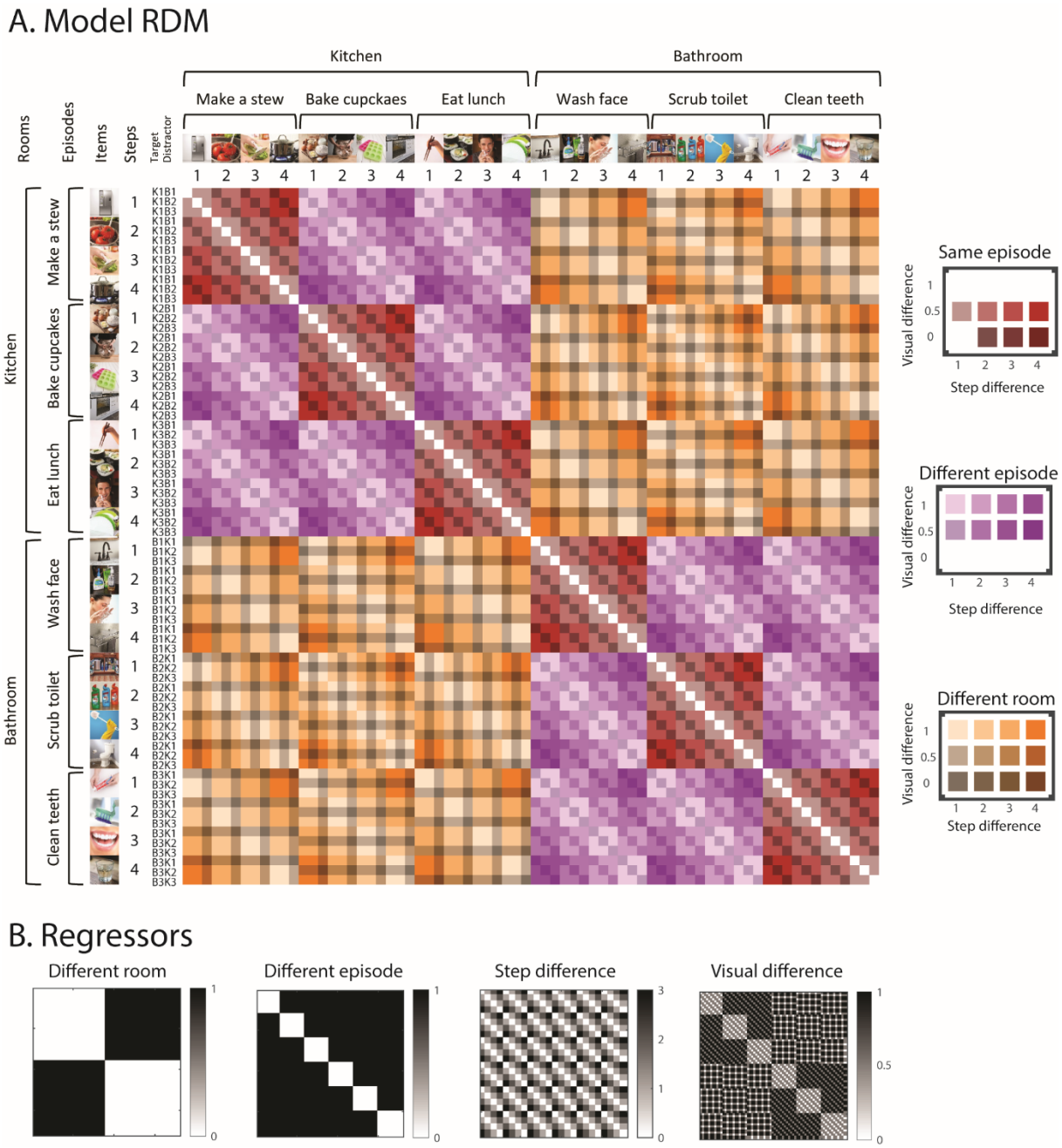


Figure B.1. Illustration of representational similarity analysis. A. Conceptual RDM. LDC dissimilarities were computed between every possible pair of events (6 cued tasks × 4 steps × 3 distractor combinations per episode), generating a 72 × 72 RDM. Diagonal cells of the RDM are zero by definition as they do not reflect a dissimilarity between different events. Off-diagonal cells reflect pattern dissimilarity between events that differ in room, episode, step, item, and/or visual difference. These included event pairs that shared the same episode (red cells); events that shared the same room but different episodes (purple cells); and events that

differed in both episodes and rooms (orange cells). All event pairs additionally differed in item. Saturation is used to indicate the difference in steps between event pairs, and brightness is used to indicate the difference in the possible stimuli presented in the visual search arrays. The visual stimuli in each search array contains a target and distractor drawn from the cued task, and two distractors drawn from a distractor task from a different room. The six tasks are labeled K1, K2, K3, B1, B2, and B3 (with K indicating a kitchen task and B indicating a bathroom task). B. Regressors used to quantify the influence of room, episode, step, and visual difference.

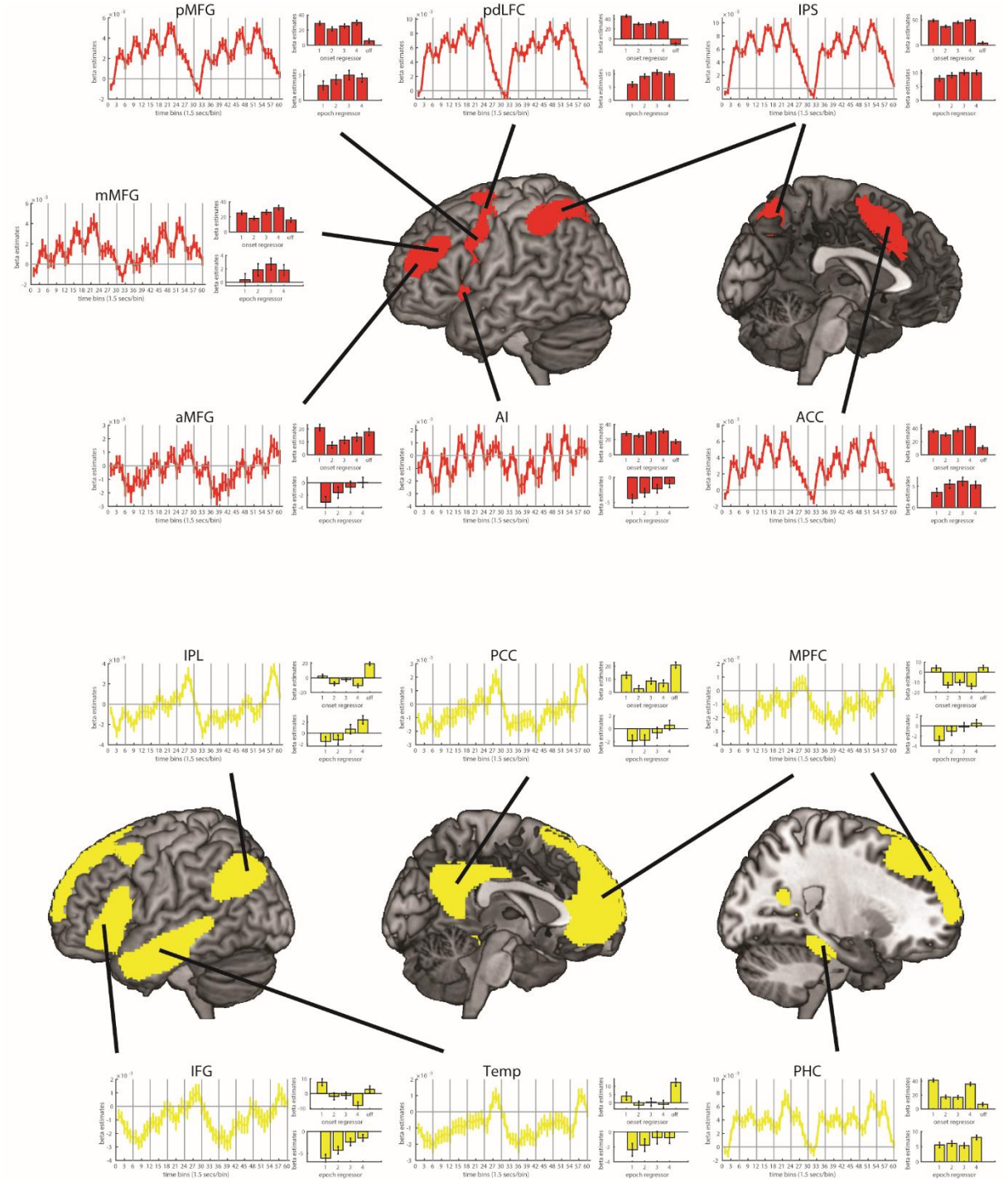


Figure B.2. FIR timecourses and activation profiles of onset and epoch responses in individual ROIs in the MD (red) and DMN (yellow) networks. The layout is the same as Figure 3.4.

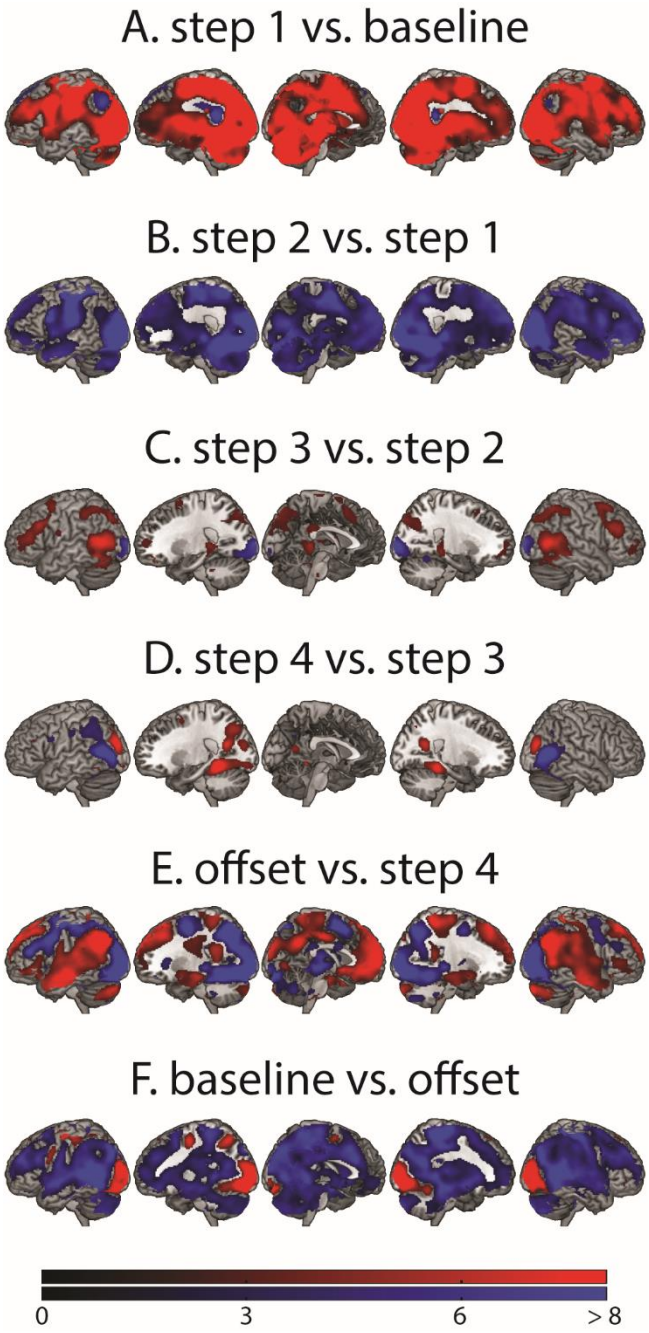


Figure B.3. Pairwise comparisons of onset/offset regressors for each transition across the episode, from A. baseline to step 1, B. step 1 to step 2, C. step 2 to step 3, D. step 3 to step 4, E. step 4 to episode offset, and F. episode offset to baseline. Positive contrasts (red) were regions that increased activation after a transition, while negative contrasts (blue) were regions that deactivated after transition. Colors indicate t-values. All activation maps are thresholded at FDR < 0.05.

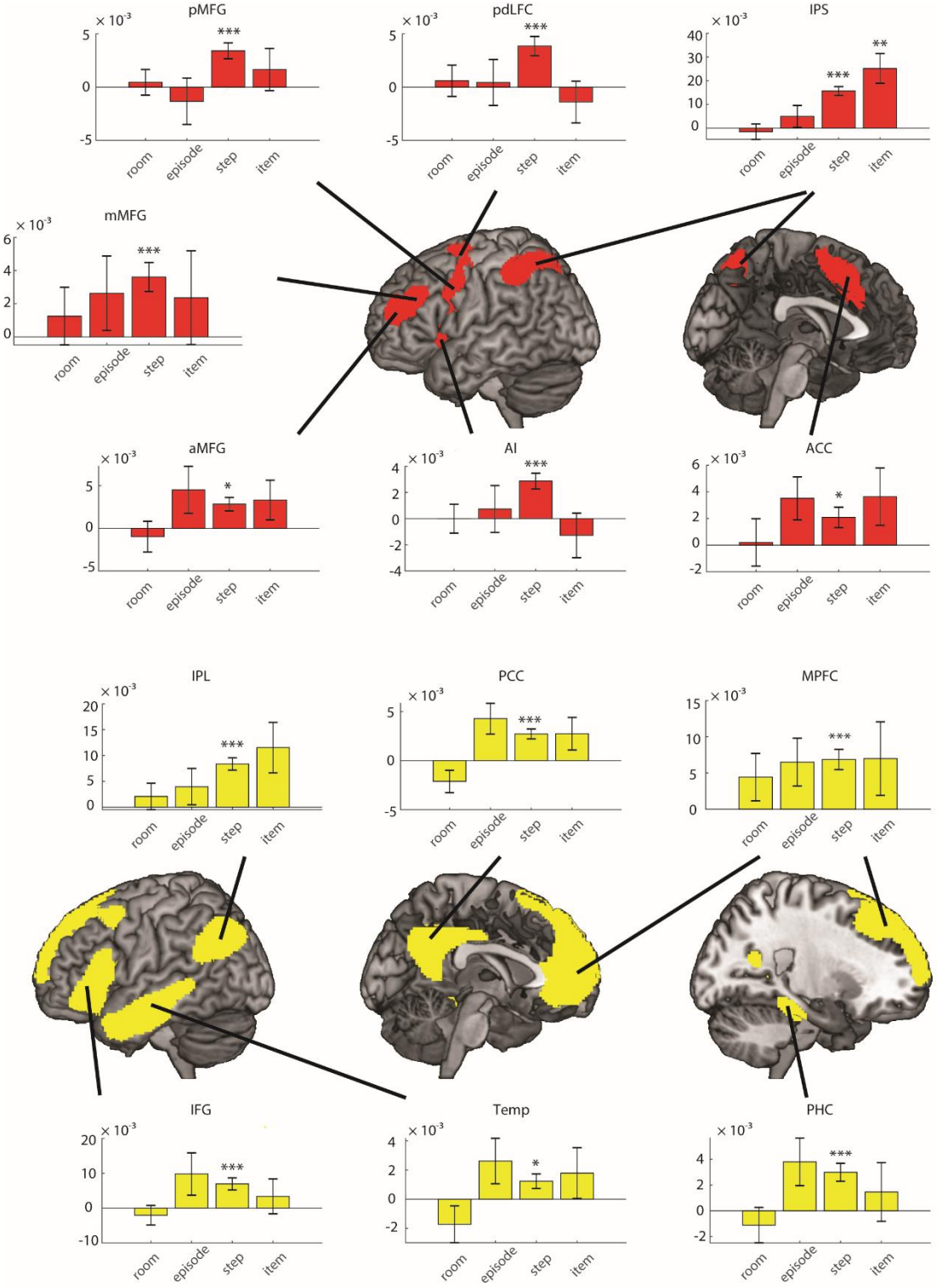


*Figure B.4. Coding of room, episode, step, and item information in individual ROIs in the MD (red) and DMN (yellow) networks. Results were FDR corrected across the number of ROIs separately for each information type. Error bars represent standard error. *** indicates $p < 0.001$, ** indicates $p < 0.01$, and * indicates $p < 0.05$ for 1-tailed t-tests against zero.*

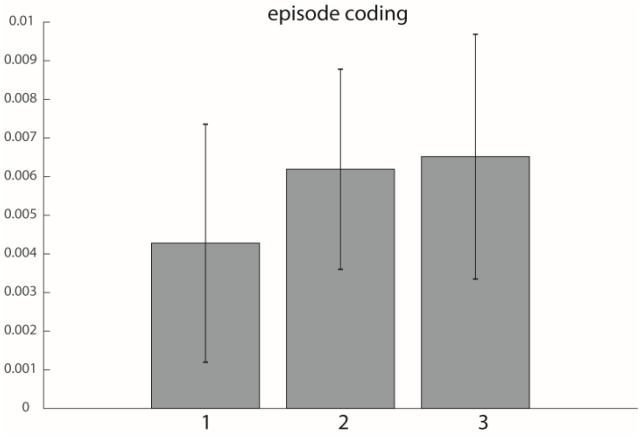


Figure B.5. Episode coding as a function of step difference in the DMN.